

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6968516号

(P6968516)

(45) 発行日 令和3年11月17日 (2021. 11. 17)

(24) 登録日 令和3年10月29日 (2021. 10. 29)

(51) Int. Cl. F I
HO 4 N 21/2662 (2011. 01) HO 4 N 21/2662
HO 4 N 21/845 (2011. 01) HO 4 N 21/845

請求項の数 15 (全 50 頁)

(21) 出願番号	特願2015-553102 (P2015-553102)	(73) 特許権者	000001007
(86) (22) 出願日	平成26年1月17日 (2014. 1. 17)		キヤノン株式会社
(65) 公表番号	特表2016-509418 (P2016-509418A)		東京都大田区下丸子3丁目30番2号
(43) 公表日	平成28年3月24日 (2016. 3. 24)	(74) 代理人	100126240
(86) 国際出願番号	PCT/EP2014/050952		弁理士 阿部 琢磨
(87) 国際公開番号	W02014/111547	(74) 代理人	100124442
(87) 国際公開日	平成26年7月24日 (2014. 7. 24)		弁理士 黒岩 創吾
審査請求日	平成29年1月6日 (2017. 1. 6)	(72) 発明者	フレデリック マゼ
審判番号	不服2019-13196 (P2019-13196/J1)		フランス国 レンヌーアタラント, セデ
審判請求日	令和1年10月2日 (2019. 10. 2)		ックス セッソナーセヴィニエ 3551
(31) 優先権主張番号	1300953.5		7, リュ ドゥ ラ トゥッシューランベ
(32) 優先日	平成25年1月18日 (2013. 1. 18)		ール キヤノン リサーチセンター フラ
(33) 優先権主張国・地域又は機関	英国 (GB)		ンス エス, エー, エス, 内

最終頁に続く

(54) 【発明の名称】 生成装置、生成方法、提供装置、提供方法、及び、プログラム

(57) 【特許請求の範囲】

【請求項 1】

複数の画像がタイル領域に分割して階層符号化されることで得られる画像データに基づいて H E V C ビットストリームをカプセル化したメディアファイルを生成する生成方法であって、

前記複数の画像のうちの第1の画像における第1のタイル領域の画像データと前記複数の画像のうちの第2の画像における前記第1のタイル領域の画像データとを格納した第1のタイルトラックと、前記複数の画像のうちの第1の画像における第2のタイル領域の画像データと前記複数の画像のうちの第2の画像における前記第2のタイル領域の画像データとを格納した第2のタイルトラックとをそれぞれ別のタイルトラックとして生成するタイルトラック生成ステップと、

初期設定データを含む初期設定データトラックを生成する初期設定データトラック生成ステップと、

前記タイルトラック及び前記初期設定データトラックとは異なるトラックを生成するステップであって、前記タイルトラック生成ステップにより生成される各タイルトラックから画像データを抽出するためのエクストラクタに、前記各タイルトラックにおけるデータオフセットを表す情報とデータ長を示す情報とを含む参照トラックを生成する参照トラック生成ステップと、

前記タイルトラック生成ステップにより生成された複数のタイルトラックと、前記参照トラック生成ステップにより生成された参照トラックとを有するメディアファイルを生成

10

20

するファイル生成ステップと、を有することを特徴とする生成方法。

【請求項 2】

前記タイルトラック生成ステップにおいては、複数のタイル領域の中から選択されたタイル領域の画像データを格納したタイルトラックが生成されることを特徴とする請求項 1 に記載の生成方法。

【請求項 3】

複数のタイル領域のうち、1 以上のタイル領域を指定する指定情報を受け付ける受付ステップを有し、

前記ファイル生成ステップにおいては、前記複数のタイル領域のうち指定情報により指定されたタイル領域の画像データを格納するタイルトラックを有し、前記複数のタイル領域のうち前記指定情報により指定されないタイル領域の画像データを格納するタイルトラックを有しないメディアファイルが生成されることを特徴とする請求項 1 に記載の生成方法。

10

【請求項 4】

前記メディアファイルは、ISO ベースメディアファイルフォーマットの規格において規定される MOV (ムービーボックス) 及び MAT (メディアデータボックス) を含む複数のボックスを有するファイルであることを特徴とする請求項 1 乃至 3 のうち、何れか 1 項に記載の生成方法。

【請求項 5】

前記メディアファイルは、時間分割された動画データを別個に格納する複数のメディアセグメントを含むことを特徴とする請求項 1 乃至 3 のうち、何れか 1 項に記載の生成方法。

20

【請求項 6】

前記ファイル生成ステップにおいて生成されたメディアファイルを、通信路を介してデバイスに提供する提供ステップを有することを特徴とする請求項 1 乃至 5 のうち、何れか 1 項に記載の生成方法。

【請求項 7】

前記タイルトラックは、前記複数の画像において同じ位置に存在するタイル領域の画像データを格納することを特徴とする請求項 1 乃至 6 のうち、何れか 1 項に記載の生成方法。

30

【請求項 8】

複数の画像がタイル領域に分割して階層符号化されることで得られる H E V C ビットストリームの提供方法であって、

前記複数の画像のうちの第 1 の画像における第 1 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 1 のタイル領域の画像データとを格納した第 1 のタイルトラックと、前記複数の画像のうちの第 1 の画像における第 2 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 2 のタイル領域の画像データとを格納した第 2 のタイルトラックとをそれぞれ別のタイルトラックとして生成するタイルトラック生成ステップと、

初期設定データを含む初期設定データトラックを生成する初期設定データトラック生成ステップと、

40

前記タイルトラック及び前記初期設定データトラックとは異なるトラックを生成するステップであって、前記タイルトラック生成ステップにより生成される各タイルトラックから画像データを抽出するためのエクストラクタに、前記各タイルトラックにおけるデータオフセットを表す情報とデータ長を示す情報とを含む参照トラックを生成する参照トラック生成ステップと、

前記タイルトラック生成ステップにより生成された複数のタイルトラックと、前記参照トラック生成ステップにより生成された参照トラックとをそれぞれ提供する提供ステップと、を有することを特徴とする提供方法。

【請求項 9】

50

前記タイルトラック生成ステップにおいては、複数のタイル領域の中から選択されたタイル領域の画像データを格納したタイルトラックが生成されることを特徴とする請求項 8 に記載の提供方法。

【請求項 10】

複数の画像がタイル領域に分割して階層符号化されることで得られる画像データに基づいて H E V C ビットストリームをカプセル化したメディアファイルを生成する生成装置であって、

前記複数の画像のうちの第 1 の画像における第 1 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 1 のタイル領域の画像データとを格納した第 1 のタイルトラックと、前記複数の画像のうちの第 1 の画像における第 2 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 2 のタイル領域の画像データとを格納した第 2 のタイルトラックとをそれぞれ別のタイルトラックとして生成するタイルトラック生成手段と、

初期設定データを含む初期設定データトラックを生成する初期設定データトラック生成手段と、

前記タイルトラック及び前記初期設定データトラックとは異なるトラックを生成する手段であって、前記タイルトラック生成手段により生成される各タイルトラックから画像データを抽出するためのエクストラクタに、前記各タイルトラックにおけるデータオフセットを表す情報とデータ長を示す情報とを含む参照トラックを生成する参照トラック生成手段と、

前記タイルトラック生成手段により生成された複数のタイルトラックと、前記参照トラック生成手段により生成された参照トラックとを有するメディアファイルを生成するファイル生成手段と、を有することを特徴とする生成装置。

【請求項 11】

前記タイルトラック生成手段は、複数のタイル領域の中から選択されたタイル領域の画像データを格納したタイルトラックを生成することを特徴とする請求項 10 に記載の生成装置。

【請求項 12】

複数の画像がタイル領域に分割して階層符号化されることで得られる H E V C ビットストリームの提供装置であって、

前記複数の画像のうちの第 1 の画像における第 1 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 1 のタイル領域の画像データとを格納した第 1 のタイルトラックと、前記複数の画像のうちの第 1 の画像における第 2 のタイル領域の画像データと前記複数の画像のうちの第 2 の画像における前記第 2 のタイル領域の画像データとを格納した第 2 のタイルトラックとをそれぞれ別のタイルトラックとして生成するタイルトラック生成手段と、

初期設定データを含む初期設定データトラックを生成する初期設定データトラック生成手段と、

前記タイルトラック及び前記初期設定データトラックとは異なるトラックを生成する手段であって、前記タイルトラック生成手段により生成される各タイルトラックから画像データを抽出するためのエクストラクタに、前記各タイルトラックにおけるデータオフセットを表す情報とデータ長を示す情報とを含む参照トラックを生成する参照トラック生成手段と、

前記タイルトラック生成手段により生成された複数のタイルトラックと、前記参照トラック生成手段により生成された参照トラックとをそれぞれ提供する提供手段と、を有することを特徴とする提供装置。

【請求項 13】

前記タイルトラック生成手段は、複数のタイル領域の中から選択されたタイル領域の画像データを格納したタイルトラックを生成することを特徴とする請求項 12 に記載の提供装置。

【請求項 1 4】

コンピュータに請求項 1 乃至 7 のうち何れか 1 項に記載の生成方法を実行させるためのプログラム。

【請求項 1 5】

コンピュータに請求項 8 又は 9 に記載の提供方法を実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、一般的には、例えば、MPEG 標準化機構により定義されたベース・メディア・ファイル・フォーマットによる、ストリーム配信を改良するための時限メディアデータのカプセル化の分野に関し、特に、圧縮されたビデオストリーム中のユーザーが選択した関心領域の HTTP (ハイパーテキスト・トランスファー・プロトコル) ストリーミングに関する。より詳しくは、本発明は、データ、特に 1 つまたは複数のタイルの効率的なストリーミングを可能にする空間タイルなどの分割されたデータを含むエレメンタリストリームをカプセル化するための方法、デバイス、およびコンピュータプログラムに関する。

10

【背景技術】

【0002】

ビデオ符号化は、ビデオ画像を送信または記憶できるように、一連のビデオ画像をコンパクトなデジタル化されたビットストリームに変換する方法である。符号化デバイスはビデオ画像を符号化するために用いられる。表示および視聴のためにビットストリームを再構築するために、関係する復号化デバイスが利用できる。一般的な目的は、当初のビデオ情報より小さいサイズになるようにビットストリームを形成することである。これによって、ビットストリームコードを送信または記憶するために必要となる転送ネットワークまたは記憶装置の容量を有利に低減する。送信するために、ビデオビットストリームは、通常ヘッダーおよびチェックビットを追加する伝送プロトコルに従って、一般にカプセル化される。

20

【0003】

近年では、ムービング・ピクチャー・エクスパート・グループ (MPEG) は、HTTP (ハイパーテキスト・トランスファー・プロトコル) 上の既存のストリーミングソリューションを統一し、取って代わるための新規な標準規格を発表した。この新規な標準規格は、「HTTP 上のダイナミック・アダプティブ・ストリーミング (DASH)」と呼ばれ、標準ウェブサーバーに基づいて HTTP 上のメディアストリーミングモデルをサポートすることを目的とし、そこでは、情報 (すなわちストリームするメディアデータの選択、ならびにユーザー選択、ネットワーク条件、およびクライアントの能力に対するビットストリームの動的な適応) がクライアント選択およびデバイスだけに依存する。

30

【0004】

本モデルでは、メディアプレゼンテーションは、データセグメントにおいて、および提供される時限メディアデータの構成を示す「メディアプレゼンテーション記述 (MPD)」と呼ばれるマニフェストにおいて構成される。具体的には、マニフェストは、データセグメントをダウンロードするために用いるリソース識別子を含み、有効なメディアプレゼンテーションを取得するためにそれらのデータセグメントを選択し結合するためのコンテキストを提供する。リソース識別子は、通常、おそらくバイト範囲と組み合わせられた HTTP-URL (ユニフォーム・リソース・ロケータ) である。マニフェストに基づいて、クライアントデバイスは、いかなる時にも、そのニーズ、その能力 (例えばサポートされたコーデック、ディスプレイサイズ、フレームレート、品質レベルなど) に従って、およびネットワーク条件 (例えば利用可能なバンド幅) に応じて、どのメディアセグメントをメディアデータサーバーからダウンロードするべきかを決定する。

40

【0005】

さらに、ビデオ解像度は絶えず増加しており、標準精細度 (SD) から高精細度 (HD

50

）に、さらに超高精細度（例えば4 K 2 Kまたは8 K 4 K、すなわち4,096×2,400画素または7,680×4,320画素の画像を含むビデオ）へと進んでいる。しかしながら、全ての受信およびビデオ復号化デバイスが、フル解像度のビデオにアクセスするためのリソース（例えばネットワークアクセスバンド幅またはCPU（中央処理ユニット））を備えているわけではなく、特に、ビデオが超高精細度である場合には、全てのユーザーがこのようなビデオにアクセスすることを必要とするわけではない。このような状況では、いくつかの関心領域（ROI）にのみアクセスする、すなわちビデオシーケンス全体のいくつかの空間的下位部分にのみアクセスする能力を提供することが、特に有益である。

【0006】

ビデオに属するフレームの空間的下位部分にアクセスする既知の機構は、一般にタイルと呼ばれる独立にデコード可能な空間領域の配置としてビデオの各フレームを構成することにある。SVC（スケーラブルビデオ符号化）またはHEVC（高効率ビデオ符号化）などのいくつかのビデオフォーマットは、タイル定義のためのサポートを提供する。ユーザーが定義するROIは、1つあるいは複数の隣接するタイルをカバーすることができる。

【0007】

したがって、HTTPプロトコルに従ってユーザーが選択したROIをストリーミングするために、1つまたは複数のタイルに空間的にアクセスすることを可能にし、アクセスされたタイルの組み合わせを可能にする方法で、符号化されたビデオビットストリームの時限メディアデータのカプセル化を提供することが重要である。

【0008】

符号化されたビデオビットストリームは、完全なフレームに対応する1組の隣接する時間的サンプルとして一般に構築され、時間的サンプルは復号化する順序の関数として構成されることを想起すべきである。このような符号化されたビットストリームをカプセル化し記述するために、ファイルフォーマットが用いられる。

【0009】

説明のために、国際標準化機構のベース・メディア・ファイル・フォーマット（ISOBMFF）は、よく知られたフレキシブルで拡張可能なフォーマットであり、それは、ローカル記憶装置、またはネットワークもしくは別のビットストリーム配信機構を介した伝送のために符号化された時限メディアデータ・ビットストリームを記述する。このファイルフォーマットは、オブジェクト指向である。それはボックスと呼ばれるビルディングブロックから構成され、ビルディングブロックはシーケンシャルに、または階層的に構成され、タイミングおよび構造パラメータなどの、符号化された時限メディアデータ・ビットストリームのパラメータを定義する。このファイルフォーマットによれば、時限メディアデータ・ビットストリームは、トラックボックスと呼ばれる別のデータ構造で定義されるmdatボックスと呼ばれるデータ構造に収容される。トラックは、サンプルの時限シーケンスを表しており、そこでは、サンプルが単一のタイムスタンプと関係する全てのデータに対応する、すなわち、単一フレームと関係する全てのデータ、または同じタイムスタンプを共有するいくつかのフレームと関係する全てのデータに対応している。

【0010】

SVCフォーマットのビデオなどのスケーラブルなビデオでは、階層化されたメディアデータ構成を複数の従属するトラックを用いて効率的に表現することができて、各トラックがスケーラビリティの特定のレベルでビデオを表現する。トラック間のデータ重複を回避するために、エクストラクタを用いることができる。標準ファイルフォーマットによれば、エクストラクタは、ビットストリームに直接含まれるデータ構造であって、他のビットストリームからネットワークアブストラクション層（NAL）ユニットの効率的な抽出を可能にする。例えば、拡張層トラックのビットストリームは、ベース層トラックからNALユニットを参照するエクストラクタを含むことができる。その後、このような拡張層トラックがファイルフォーマットから抽出された場合に、エクストラクタはそれらが参照

10

20

30

40

50

しているデータと置き換えられなければならない。

【0011】

これらの機構を埋め込んでいるISO BMFFを用いる場合に、下位情報を記述するために、およびこの下位情報へのアクセスを容易にするために、または効率的にビットストリームを複数セグメントに構成するために、いくつかの戦略を採用することができる。

【0012】

例えば、「Implications of the ISO Base Media File Format on Adaptive HTTP Streaming of H.264/SVC」というタイトルの論文において、著者であるKofler他は、ISO BMFFの可能性ならびに限界を考慮したHTTPストリーミングのためのスケラブルなビデオビットストリーム(H.264/SVC)を構成するための3つの異なる戦略を以下のように提示している。

【0013】

a) 単一ファイルは、全てのISO BMFFメタデータ(トラック定義を含む)を含むファイルタイプボックス「ftyp」およびムービーボックス「moov」を含む特定のファイルヘッダーを備えており、また単一ファイルは、符号化されたビットストリーム全体を含む単一のmdatボックスも備えている。この構成は、ローカル記憶装置には好適であるが、クライアントがビットストリーム全体の一部のみを必要とする場合があり得るHTTPストリーミングには適さない。

【0014】

b) 単一ファイルは、フラグメンテーションのために好適な複数のmoof/mdatボックスを備えている。このフォーマットは、進歩的なダウンロードを可能にする。moofボックスは、フラグメントレベルではmoovボックスと等価である。この方式によれば、断片化したメディアファイルを用いて、スケラブルなビットストリームは、異なるスケラビリティレベルでビデオを表現する複数の従属するトラックに分割される。エクストラクタは、他のトラックからNALユニットを参照するために用いられる。タイルごとに1つのトラックが用いられる場合には、全てのアドレス可能なトラックを前もって準備しなければならない。トラックは独立に選択することができない。いくつかのタイルを表示しようとする場合には、いくつかのビットストリームを復号化しなければならない。ベース層は数回復号化される。

【0015】

c) 複数セグメントファイルは、各ファイルがそれ自体のURLによってアクセス可能であって、独立にダウンロード可能である。各セグメントは、通常、一種のファイルヘッダーとして作用するセグメントタイプボックス(styp)、任意のセグメントインデックスボックス(sidx)、および1つもしくは複数のフラグメントから構成される。更にまた、各フラグメントは、moofおよびmdatボックスから構成される。この方式によれば、断片化したメディアファイルを用いて、各トラックは、スケラビリティの1つのレベルに関して関係するビットストリームを用いてそれ自体のセグメントに記憶される。必要に応じて、エクストラクタは、従属するトラックから必要なビットストリームを参照するために用いられる。このような符号化方式は、トラックを独立にストリーミングするのに特に適している。それはDASH標準規格によく適合しているが、いくつかのビットストリームを復号化しなければならない。したがって、トラックごとに1つのデコーダが必要となるので、それはタイルストリーミングには適さない。さらに、2つ以上のタイルを選択する場合に、ベース層のビットストリームが重複する可能性がある。

【0016】

空間タイルに適用される場合には、これらの戦略のいずれもHTTPストリーミングのコンテキストにおいて特定のタイルに効果的にアクセスすることができない。実際、既存のファイルフォーマット定義では、符号化されたビットストリームの多数の非連続バイト範囲にアクセスすることが依然として必要であり、あるいは、所与の時間間隔に対応するいくつかのフレームの空間タイルを表示するために、ビットストリームの重複が生じる。

10

20

30

40

50

【 0 0 1 7 】

これらの問題を解決するために、空間タイルに好適な効率的なデータ構成およびトラック記述方式が提供され、それは、どのようなトラックの組み合わせがクライアントアプリケーションによって選択されても、ISO BMFF パーシングの結果が常にビデオデコーダのための有効なビデオ・エレメンタリ・ビットストリームをもたらすことを保証する。

【 発明の概要 】

【 0 0 1 8 】

本発明の生成方法は、複数の画像がタイル領域に分割して階層符号化されることで得られる画像データに基づいてH E V Cビットストリームをカプセル化したメディアファイルを生成する生成方法であって、前記複数の画像のうちの第1の画像における第1のタイル領域の画像データと前記複数の画像のうちの第2の画像における前記第1のタイル領域の画像データとを格納した第1のタイルトラックと、前記複数の画像のうちの第1の画像における第2のタイル領域の画像データと前記複数の画像のうちの第2の画像における前記第2のタイル領域の画像データとを格納した第2のタイルトラックとをそれぞれ別のタイルトラックとして生成するタイルトラック生成ステップと、初期設定データを含む初期設定データトラックを生成する初期設定データトラック生成ステップと、前記タイルトラック及び前記初期設定データトラックとは異なるトラックを生成するステップであって、前記タイルトラック生成ステップにより生成される各タイルトラックから画像データを抽出するためのエクストラクタに、前記各タイルトラックにおけるデータオフセットを表す情報とデータ長を示す情報とを含む参照トラックを生成する参照トラック生成ステップと、前記タイルトラック生成ステップにより生成された複数のタイルトラックと、前記参照トラック生成ステップにより生成された参照トラックとを有するメディアファイルを生成するファイル生成ステップとを有することを特徴とする。

【 図面の簡単な説明 】

【 0 0 1 9 】

本発明のさらなる利点は、図面および詳細な説明を考察することで、当業者にとって明らかになる。いかなる付加的な利点も、本明細書に組み込まれるものとする。

【 0 0 2 0 】

ここで本発明の実施形態について、単なる例として、以下の図面を参照して説明する。

【 図 1 a 】 符号化ビデオビットストリームにおいてタイルを符号化する例を示す図である。

【 図 1 b 】 符号化ビデオビットストリームにおいてタイルを符号化する例を示す図である。

【 図 2 】 ユーザーによって表示するように選択されたタイルの時間的パイプを示す図である。

【 図 3 】 特定の実施形態による、スケーラブルなメディアデータトラックを用いる初期化セグメントファイルのブロック図の例を示す。

【 図 4 】 特定の実施形態による、タイルトラックおよび1つの複合トラックを含むメディアセグメントファイルのブロック図の例を示す図であり、図3に示したような初期化セグメントファイルにおいて宣言されたトラックに対応する。

【 図 5 】 所与の時間的期間の連続するビデオフレームの空間部分を表現する有効な復号可能な時限メディアデータ・ビットストリームを構築するために、ダウンロードされたメディアデータセグメントを連結する例を示す図である。

【 図 6 a 】 特定の実施形態による、サーバーとクライアントデバイスとの間で時限メディアデータを送信するためのステップを示すフローチャートである。

【 図 6 b 】 特定の実施形態による、サーバーとクライアントデバイスとの間で時限メディアデータを送信するためのステップを示すフローチャートである。

【 図 7 】 クライアントデバイスによって受け取られた連結されたメディアセグメントからの有効な時限メディアデータ・ビットストリームの生成を示すフローチャートである。

【図 8】1つまたは複数の実施形態のステップを実行することができる、サーバーまたはクライアントデバイスのブロック図である。

【図 9 a】H E V C ビットストリームにおけるタイルおよびスライスセグメントの例を示す図である。

【図 9 b】H E V C ビットストリームにおけるタイルおよびスライスセグメントの例を示す図である。

【図 9 c】H E V C ビットストリームにおけるタイルおよびスライスセグメントの例を示す図である。

【図 1 0】本発明の一実施形態による、複合トラックおよび独立タイルトラックを含む 1 組のトラックとして H E V C ビットストリームをカプセル化する例を示す図である。

10

【図 1 1】H E V C ビットストリームをカプセル化するための、第 1 の実施形態による、メディアデータトラックを用いる初期化セグメントファイルのブロック図の例を示す。

【図 1 2 a】フルビデオにおけるタイルの位置、タイルのサイズ、および様々なタイル化構成を扱うのに適したタイルトラックがサブサンプルレベルでアーチファクトなしに復号化できることの通知を示す図である。

【図 1 2 b】フルビデオにおけるタイルの位置、タイルのサイズ、および様々なタイル化構成を扱うのに適したタイルトラックがサブサンプルレベルでアーチファクトなしに復号化できることの通知を示す図である。

【図 1 3】本発明の一実施形態による、標準ビデオトラックとして再生可能な複合トラックおよび独立タイルトラックを含む 1 組のトラックとして H E V C ビットストリームをカ

20

プセル化する例を示す図である。

【図 1 4】本発明の別の実施形態による、標準ビデオトラックとして再生可能な複合トラック、初期化データトラック、および独立タイルトラックを含む 1 組のトラックとして H E V C ビットストリームをカプセル化する例を示す図である。

【発明を実施するための形態】

【0 0 2 1】

特定の実施形態によれば、時限サンプル（例えば画像）を含むタイル化時限メディアデータ（例えばビデオデータ）などの分割された時限メディアデータは、通常はベース層トラックおよびいくつかのタイルトラックである 1 組のいくつかの時限メディアデータトラック、ならびに時限メディアデータトラックへの参照を含む参照トラックもしくは複合トラックとして送信される。各タイルトラックは、いくつかの時限サンプルの 1 つの空間サブサンプル（例えばいくつかの N A L ユニット）を含む。複合トラックから時限メディアデータトラックを参照するために、拡張されたエクストラクタタイプが定義される。時限メディアデータトラックは、表示可能でないとラベルされ、タイルのための時限メディアデータを運び、それを記述する。このような 1 組の時限メディアデータトラックおよび複合トラックは、空間ビデオタイルの選択、構成、および効率的なストリーミングを可能にする。各トラックは、1 組のメディアセグメントファイルとしてサーバーデバイスからクライアントデバイスへ送信され得る。メディアセグメントファイルを復号化するのに必要なメタデータを送信するために、初期化セグメントファイルを用いることができる。

30

【0 0 2 2】

図 1 は、図 1 a および図 1 b を含み、符号化ビデオビットストリームにおいてタイルを符号化する例を示す。

40

【0 0 2 3】

例示のために、以下の説明では、各ビデオフレーム（時限サンプル）がビデオフレームの空間的下位部分（空間サブサンプル）に対応する独立に復号可能なタイルから構成されていると考える。ビデオは、好ましくはスケラブルであって、スケラビリティの種々のレベルで構成される。図 1 a に示すように、ビデオフレーム 1 0 0 は、H D ベース層（1 0 2）および 4 K 2 K 拡張層（1 0 4）を含むことができる。さらに例示のために、拡張層 1 0 4 は、a、b、c、d で示す 4 つの正規のタイルに分割することができる。種々の形状のタイルを扱うことができる点に留意する必要がある。同様に、ベース層 1 0 2 を

50

いくつかのタイルに分割することができる。このような場合、いくつかの複合トラックを、例えばベース層について1つの、および拡張層もしくは拡張層の各々について1つの複合トラックを用いることができる。

【0024】

本発明がスケーラブルなビデオフォーマットに限定されない点にも留意する必要がある。それは、タイルを独立に復号化することができる全てのビデオフォーマットに当てはまる可能性がある。したがって、MPEG4、AVC、HEVC、SVC、または将来のSHVCなどの任意のビデオ圧縮アルゴリズムを、本発明の一実施形態と共に用いることができる。

【0025】

図1bは、復号化する順序で典型的な符号化ビデオビットストリームを表す。図示するように、符号化ビデオビットストリームは、ここでは時間的順序で符号化された3つのビデオフレーム(110、112、および114)を含む。各ビデオフレームは、ベース層(BL)の全てのネットワークアブストラクション層(NAL)ユニットを含み、その後には拡張層のNALユニットが続く。例えば、第1のビデオフレーム(110)のベース層(102-1)のNALユニット(1BL、116)の後に、第1のビデオフレームの拡張層(104-1)のNALユニット(1共通、1a、1b、1c、1d、118)が続く。

【0026】

ビデオビットストリームの、空間タイルを有する拡張層に対応する部分は、各タイルのNALユニットから構成される。選択的に、それは、全てのタイルに共通で、タイルのいずれかを復号化するのに必要なNALユニットを含むこともできる。所与のフレームの全てのタイルに共通なNALユニットは、ビデオビットストリームの対応する部分のどこにでも(すなわちビデオフレームのタイルのNALユニットの前、それらの間、またはそれらの後に)位置することができる。

【0027】

図示するように、ビデオビットストリームの、空間タイルb、c、およびdを含む第1のビデオフレーム(110)の拡張層に対応する部分は、タイルごとのNALユニット(1a、1b、1cおよび1d)および全てのタイルa、b、c、およびdに共通なNALユニット(1共通)から構成される。

【0028】

図2は、ユーザーによって表示するように選択されたタイルの時間的パイプを示す。より正確には、図2は、第1のビデオフレームnおよび第2のビデオフレームn+m(ここでnおよびmは整数値である)を表し、第1および第2のビデオフレームの各々は1から12までの番号を付した12のタイルを含む。これらの12のタイルのうち、第3および第7のものだけが表示される(太線で示す)。ビデオフレームnおよびn+mは、所与の時間的期間に対応する一連の連続フレームに属する。したがって、フレームnからフレームn+mまでの各フレームの第3および第7のタイルは、連続的に表示される。

【0029】

図1に示したように、ビデオビットストリームのデータは、フルフレームに対応する時間的サンプルとして構成される。したがって、図2を参照して上述したように、これらのフレームの特定の空間的領域が所与の期間の間にアクセスされる場合には、フレームごとにいくつかの小さいバイト範囲にアクセスする必要がある。これは、生成された要求の数に関して、およびデータオーバーヘッドに関して、HTTPストリーミングにおいて非効率的である。

【0030】

したがって、ROIストリーミングのための圧縮ビデオのより効果的なアクセスを提供するためには、特定のタイルのデータが所与の期間(すなわち1組の連続フレーム)について隣接するバイト範囲(パイプを形成する)として構成されるように、時限メディアデータ・ビットストリームを再構成すべきである。

10

20

30

40

50

【0031】

したがって、ビデオフレームの空間的下位部分のみを表示する場合には、選択された空間領域に対応するタイルのパイプのみを、パイプごとに、および期間ごとに1つのHTTP要求を用いてダウンロードしなければならない(例えば図2のタイル3および7)。

【0032】

図3は、特定の実施形態による、メディアデータトラックを用いる初期化セグメントファイルのブロック図の例を示す。

【0033】

図3に示す初期化セグメントファイルおよび図4に示すメディアセグメントファイルは、いくつかのトラックを独立にストリーミングすると共に、ビデオビットストリームをISOベース・メディア・ファイル・フォーマットに再構成しカプセル化するための本発明の一実施形態を示す。例示のために、図3および図4のコンテンツは、図1に示したビデオビットストリームに基づいている。

【0034】

いくつかのトラックを独立にストリーミングすると共に、ビデオビットストリームをISO BMFFに再構成しカプセル化するために、タイルトラックと呼ばれる新しいタイプのトラックを定義する。タイルトラックは、定義上、サンプルが単一のタイムスタンプと関係する全てのデータを表現する、関連したサンプルの時限シーケンスであるトラックである。サンプルが通常は個々のビデオフレームである既知のビデオメディアトラックとは逆に、タイルトラックのサンプルは、フルビデオフレームの空間的に定義された下位部分を定義する。したがって、タイルトラックは、所与のタイルに関連したNALユニットだけを含む。このようにして、独立なセグメントファイルの各トラックを記憶することによって、隣接するバイト範囲を有するタイルのパイプを生成することができる。

【0035】

しかし、タイルトラックが復号化し再生するのに必要な全てのNALユニットおよび情報を含むわけではないので、「複合トラック」と呼ばれる付加的なトラックが用いられる。特定の実施形態によれば、複合トラックは、完全なタイル化フレーム(すなわち全てのタイルの組成)を表すメディアトラックである。複合トラックは、図1を参照して説明したように(復号化する順序でフレームごとに)、その典型的構成に符号化された時限メディアデータ・ビットストリームの表現である。複合トラックは、それらのそれぞれのタイルトラックのNALユニットを参照するために、エクストラクタオブジェクトを用いる。さらに、それは、このようなNALユニットが存在する場合には、全てのタイルに共通なNALユニットを含むことができる。

【0036】

特定の実施形態によれば、他のメディアセグメントファイルにカプセル化される時限メディアデータ・ビットストリームを定義するのに必要な全てのメタデータを送信するために、初期化セグメントファイルを用いる。図3に示すように、初期化セグメントファイル300は、ファイルタイプボックス「ftyp」302およびムービーボックス「moov」304を含む。ファイルタイプボックス302は、好ましくは、セグメントファイルがどのISO BMFF仕様に対応するかを識別して、その仕様のバージョン番号を示す。ムービーボックス「moov」304は、メディアセグメントファイルに記憶されたプレゼンテーションを記述する全てのメタデータ、および特にプレゼンテーションで利用可能な全てのトラックを提供する。

【0037】

ムービーボックス「moov」304は、図1で例として提供されたスケーラブルなビデオビットストリームに対応するトラック(「トラック」ボックス306-1~306-6)の各々のための定義を含む。

【0038】

トラックボックス306-1は、ベース層(トラックID=1)を表し、4つのトラックボックス306-2~306-5(トラックボックス306-3および306-4は図

10

20

30

40

50

示せず)は、拡張層(トラックID = 2 ~ 5)の4つのタイルa、b、c、およびdを表し、トラックボックス306-6は、拡張層(トラックID = 6)を記述する複合トラックを表す。

【0039】

各トラックボックスは、少なくとも、一般的に符号308で示すトラックヘッダーボックス「tkhd」および一般的に符号310で示すトラックメディアボックス「mdia」を含む。トラックが他のトラックからのデータに依存する場合には、トラック参照ボックス「trf」もある。図示するように、識別子トラックID = 6を有する複合トラックは、トラックが識別子トラックID = 1 ~ 6を有するトラックからのデータに依存することを示すトラック参照ボックス「trf」312を含む。

10

【0040】

時限メディアデータ・ビットストリームをカプセル化するために用いるISO BMF仕様に応じて、他のボックスが必須であってもよいし、あるいは任意であってもよい点に留意する必要がある。しかし、本発明の実施形態は、適用可能なこれらのボックスに依存しないので、それらについてはここで提示しない。

【0041】

トラックヘッダーボックス「tkhd」308は、トラックの特性を指定する。情報のいくつかの項目の中で、それは、トラックの識別子(トラックID)、トラックの継続時間、および/またはトラックのビジュアルプレゼンテーションサイズ(すなわち表示エリアの幅および高さ)を提供する。またそれは、トラックが再生可能か否かを示すフラグパラメータを含む。

20

【0042】

一実施形態によれば、タイルトラックのトラックヘッダーフラグのデフォルト値は0(トラックイネーブル = 0、トラックインムービー = 0、トラックインプレビュー = 0)であって、タイルトラックがクライアントデバイスによるローカル再生およびプレビューについて無視されることを意味する。別の実施形態では、トラックがタイルトラックであることを通知するために、新しいトラックヘッダーフラグを生成することができる。

【0043】

トラックメディアボックス「mdia」310は、トラック内の時限メディアデータのパラメータを宣言するために用いる全てのオブジェクトを格納するコンテナであるとみなすことができる。それは、少なくとも、一般的に符号314で示すメディアヘッダーボックス「mdhd」、一般的に符号316で示すハンドラ参照ボックス「hdlr」、および一般的に符号318で示すメディア情報ボックス「minf」を含む。

30

【0044】

ハンドラ参照ボックス「hdlr」316は、トラックの時限メディアデータが提示されるプロセス、および、したがってトラックの時限メディアデータの性質を宣言する。例えば、ビデオトラックは、ビデオハンドラ(「vide」に等しいハンドラタイプの属性により示す)によって扱われる。ビデオサンプルは、ビジュアルサンプルエントリ(「VisualSampleEntry」)タイプのオブジェクトを用いて記述することができる。特定の実施形態によれば、タイルハンドラと呼ばれる新しいハンドラタイプ(「tile」に等しいハンドラタイプの属性により示す)は、トラックが空間サブサンプル情報を含むことを示すように定義される。符号化フォーマットに応じて、ビジュアルサンプルエントリ(「VisualSampleEntry」)タイプのオブジェクトがタイルトラックにサンプルを記述することができない場合には、サンプルを記述するためにタイトルサンプルエントリ(「TileSampleEntry」)タイプの特定のオブジェクトを定義することができる。

40

【0045】

メディア情報ボックス「minf」318は、トラックの時限メディアデータの特性情報を定義する全てのオブジェクトを含む。例えば、ベース層について、および複合トラックにおいて定義された拡張層については、「minf」ボックスは、標準ビデオメディア

50

・ヘッダーボックス「vmhd」(320)を含むことができる。

【0046】

タイルトラックに関して、新しいタイルハンドラに対応するタイルメディア・ヘッダーボックス(tmhd、322)と呼ばれる特定のボックスを、タイルのための、符号化から独立している一般プレゼンテーション情報を定義するために用いる。具体的には、それは複合ビデオトラックによって表されるビデオ解像度と関連して、タイルによってカバーされる空間的領域の幾何学的情報を含むことができ、それは次のように定義することができる。

```
aligned(8) class TileMediaHeaderBox
extends FullBox('tmhd', version = 0, 0)
{
    unsigned int(16) horizontal_offset;
    unsigned int(16) vertical_offset;
}
```

【0047】

上述したように、複合トラックは、プレゼンテーションの別のトラックに入力された参照を提供する特定のトラック参照ボックス「tref」312を含む。特定の実施形態によれば、このような入力された参照は、複合トラックからそれが参照するタイルトラックへリンクを確立するために用いることができる「タイル」参照(324)、および、この参照を含むトラックからそれが依存する時限メディアデータトラック(例えばベース層トラック(トラックID=1))へリンクを確立するために用いることができる「scal」参照(326)を含むことができる。

【0048】

図4は、特定の実施形態による、タイルトラックおよび1つの複合トラックを含むメディアセグメントファイルのブロック図の例を示す図であり、図3に示したような初期化セグメントファイルにおいて宣言されたトラックに対応する。上述したように、図3に示す初期化セグメントファイルおよび図4に示すメディアセグメントファイルは、いくつかのトラックを独立にストリーミングすると共に、ビデオビットストリームをISOベース・メディア・ファイル・フォーマットに再構成しカプセル化するための本発明の一実施形態を示す。

【0049】

図4に示すように、メディアセグメントファイル400-1~400-6(メディアセグメントファイル400-3~400-5は図示せず)は、DASH標準規格に規定されているように、一般的に符号402で示すセグメントタイプボックス「styp」、一般的に符号404で示す少なくとも1つのムービーフラグメントボックス「moov」、および一般的に符号406で示す少なくとも1つのメディアデータボックス「mdat」を含む。メディアセグメントファイルは、HTTP-URLと関係する。

【0050】

時限メディアデータ・ビットストリームをカプセル化するために用いるISO BMFF仕様に応じて、他のボックスが必須であってもよいし、あるいは任意であってもよい点に留意する必要がある。しかし、本発明の実施形態は、適用可能なこれらのボックスに依存しないので、それらについてはここで提示しない。

【0051】

しかし、セグメントタイプボックス「styp」402のフォーマットは、図3のファイルタイプボックス「ftyp」302の1つに類似しており、その参照はそのファイルがメディアセグメントファイルであることを示している。

【0052】

ムービーフラグメントボックス404は、ムービーボックス「moov」内に一般的に記憶される情報を提供する。そのヘッダー(「mfhd」)は、ムービーフラグメントごとに増加するシーケンス番号(図4ではseq_numで示す)を含む。このようなシー

ケンス番号によって、クライアントデバイスは、増加する順に受け取ったセグメントファイルを連結して、（必要に応じて）シーケンスの完全性を検証することができる。ムービーフラグメントボックス404は、関連するメディアデータボックス（「mdat」、406）のデータを有するトラックごとに、トラックフラグメントボックス「traf」（一般的に符号408で示す）を含む。トラックフラグメントボックス408は、一般的に符号410で示すトラックフラグメントヘッダーボックス「tfhd」を含み、それは、対応するメディアデータボックス（「mdat」、406）に存在するトラックのビットストリームの識別子（トラックID）を記憶するために用いる。

【0053】

メディアデータボックスは、時限メディアデータを一般に含む。標準ビデオトラックにおいては、それはビデオフレームを含む。タイトルトラックにおいては、メディアデータボックス406は、完全なビデオフレームの空間的に定義された下位部分を含む。例示のために、トラック識別子トラックID=2と関係するメディアデータボックスは、拡張層のタイトルに対応する全てのNALユニットを含む。

【0054】

複合トラック（図4のトラックID=6）では、メディアデータボックス406は、タイトルごとに、および従属する層ごとにエクストラクタ（図4ではEで示す）を含み、全てのタイトルに共通なNALユニットを含む（存在する場合）。

【0055】

図4に示すように、複合トラックと関係するメディアセグメントファイル400-6のメディアデータボックス406は、特に以下を含む。

【0056】

すなわち、ベース層トラックと関係するメディアセグメントファイル400-1のメディアデータボックス406に記憶されたベース層トラック内に符号化されたベース層データ（NALユニット1BL）へのリンクを提供する第1のエクストラクタ412-1、

いくつかのタイトルに共通なNALユニット412-2、

拡張層の第1のタイトルトラックと関係するメディアセグメントファイル400-2のメディアデータボックス406内に符号化された第1のタイトルの拡張層データ（NALユニット1a）へのリンクを提供する第2のエクストラクタ412-3、

拡張層の第2のタイトルトラックと関係するメディアセグメントファイル400-3（図示せず）のメディアデータボックス406内に符号化された第2のタイトルの拡張層データ（NALユニット、1b）へのリンクを提供する第3のエクストラクタ412-4、

拡張層の第3のタイトルトラックと関係するメディアセグメントファイル400-4（図示せず）のメディアデータボックス406内に符号化された第3のタイトルの拡張層データ（NALユニット、1c）へのリンクを提供する第4のエクストラクタ412-5、

および、拡張層の第4のタイトルトラックと関係するメディアセグメントファイル400-5（図示せず）のメディアデータボックス406内に符号化された第4のタイトルの拡張層データ（NALユニット、1d）へのリンクを提供する第5のエクストラクタ412-6である。

【0057】

エクストラクタ412-1のおかげで取得することができるNALユニットは、NALユニット412-2およびエクストラクタ412-3~412-6のおかげで取得することができるNALユニットを用いて、拡張層を完全に復号化することができるフレームのベース層の復号化を可能にする。図4から分かるように、フレームの空間部分のみが復号化されれば、全てのメディアセグメントファイル400-2~400-5（すなわちタイトルトラックに対応するビットストリーム）をダウンロードする必要はない。

【0058】

特定の実施形態によれば、エクストラクタは、以下のシンタックスを有するファイルフォーマット内部構造である。

```
class aligned(8) Extractor {
```

10

20

30

40

50

```

NALUnitHeader();
unsigned int(8) track__ref__index;
signed int(8) sample__offset;
unsigned int((lengthSizeMinusOne +
1) * 8)
        data__offset;
unsigned int((lengthSizeMinusOne +
1) * 8)
        data__length;
}

```

10

【0059】

ここで、NALユニットヘッダー() (NALUnitHeader()) は、ビデオビットストリームを符号化するために用いる符号化フォーマットに準拠するNALユニットの最初の4バイトを表す。これらの4バイトは、NALユニットをエクストラクタとして識別する(例えば、SVCでは、属性nalユニットタイプ(nal__unit__type)はエクストラクタNALユニットタイプ(タイプ31)に設定される)。

【0060】

トラック参照インデックス(track__ref__index)の値は、複合トラックのタイプ「scal」または「tile」のトラック参照ボックス「tref」において、データが抽出されるトラックを見いだすために用いるインデックスを指定する。サンプルオフセット(sample__offset)の値は、情報源として用いられるリンク先のトラックのサンプルの相対インデックスを与える。データオフセット(data__offset)およびデータ長(data__length)の値は、それぞれ、コピーする参照サンプル内の最初のバイトのオフセット、およびコピーするバイト数である。

20

【0061】

例示のために、および図3に関して、所与のエクストラクタのトラック参照インデックス(track__ref__index)の値が2に等しい場合には、これは、エクストラクタがtrefボックスの第2のエントリによって識別されるトラックを参照することを意味する(すなわち、識別子トラックID=2を有するトラック、すなわち、タイルaのタイルトラック、参照トラック(例えばベース層)を表す第1のインデックス)。

30

【0062】

図5は、所与の時間的期間の連続するビデオフレーム(ここでは2つの連続フレームに対応する)の空間部分を表現する有効な復号可能な時限メディアデータ・ビットストリームを構築するために、ダウンロードされたメディアデータセグメントを連結する例を示す図である。同じ図は、他の時間的期間について繰り返すことができる。

【0063】

図3および図4に関して説明したように、時限タイル化メディアデータ・ビットストリームは、好ましくは、1つの初期化セグメントファイルおよび複数のメディアセグメントファイルから成る1組のデータとして送信され、後者はいくつかのタイルトラックおよび1本の複合トラックを含む。

40

【0064】

初期化セグメントファイルは、各トラック上の一般情報、特にトラックのタイプ(例えばメディアトラック(音声もしくはビデオ)またはタイルトラック)、符号化フォーマット、フレーム解像度、およびトラック(トラック参照ボックス「tref」において与えられた)間の従属性を提供するムービーボックス(「moov」)を含む。これらのデータは、ダウンロードされたメディアセグメントファイル进行处理するために用いられる。図1、図3、および図4に関して説明した例に示すように、初期化セグメントファイルのムービーボックスのコンテンツは、特に、以下を含むことができる。

MOOV

トラック1: ベース層

50

トラック 2 : タイル a
トラック 3 : タイル b
トラック 4 : タイル c
トラック 5 : タイル d
トラック 6 : 拡張層

```
t r e f ( s c a l ) : ト ラ ッ ク I D = 1
t r e f ( t i l e ) : ト ラ ッ ク I D = 2
                     ト ラ ッ ク I D = 3
                     ト ラ ッ ク I D = 4
                     ト ラ ッ ク I D = 5
```

10

【 0 0 6 5 】

図5は、必要なメディアセグメントファイル（ここではベース層ならびに拡張層のタイトルaおよびcに対応する）だけをサーバーからダウンロードする場合に、メディアセグメントを連結することによって得られるファイルフォーマットを概略的に示す。このような機構は、必要とするメディアセグメントファイルだけのダウンロードを可能にするだけではなく、重複するデータのダウンロードを防止することもできる点に留意すべきである。

【 0 0 6 6 】

図示するように、ベース層トラック 504（スケーラビリティの場合には）から、および非再生可能なタイルトラック（506 および 508）からデータを参照することによって、ならびに、失われたデータを参照するエクストラクタを適切に扱うことによって（図 7 を参照して説明するように）、複合トラック 500 は、有効な復号化可能な時限メディアデータ・ビットストリーム 502 の構築を可能にする。

20

【 0 0 6 7 】

取得されたファイルフォーマットは、スケーラブルなファイルフォーマット定義に準拠する。例えば、クライアントデバイスは、ベース層トラックを選択することによってベース層だけを、あるいは複合トラックを選択することによって選択されたタイル a および c を有する拡張層だけを、再生するように決定することができる。またクライアントデバイスは、それが複合トラックを再生し続ける間に、後の時間的期間において異なる「タイルトラック」（すなわちメディアセグメントファイル）をダウンロードすることによって、タイルを表示させるように変更することもできる。

30

【 0 0 6 8 】

図 6 は、図 6 a および図 6 b を含み、特定の実施形態による、サーバーとクライアントデバイスとの間で時限メディアデータを送信するためのステップを示すフローチャートである。図 6 a に示すステップは、タイル化時限メディアデータ・ビットストリームからの R O I ストリーミングに適するセグメントファイルを生成することによって、メディアプレゼンテーションを準備するように、サーバー内で実行され、一方、図 6 b に示すステップはクライアントデバイス内で実行される。

【 0 0 6 9 】

第1のステップ(ステップ600)において、サーバーは、タイルと関係する全てのNALユニットを識別し、タイルごとに、所与のタイルに対応する全てのNALユニットから構成されるサブサンプルを含むタイルトラックを生成する。例えば、サーバーは、NALユニットの種々の領域との関連を識別するために、サブピクチャーレベルSEIメッセージに依存することができ、ならびに、HEVC標準化(提案JCTVC-K0128)で提案されているように、各ROIの位置およびサイズを識別するためのシーケンスレベルSEIメッセージに依存することができる。したがって、サーバーは、所与の期間についてタイルのパイプを生成することができる。

40

【 0 0 7 0 】

続くステップ（ステップ602）において、サーバーは、より低いレベルのスケラビリティトラック（存在する場合）にリンクされたエクストラクタ、全てのタイルに共通なNALユニット、および各タイルトラックにリンクされたエクストラクタを含む複合トラ

50

ックを生成する。エクストラクタおよび共通NALユニットは、エクストラクタをそれらが参照しているデータと置き換えることによって、復号化する順序で（図1に関して説明したように）完全なサンプルから構成された有効な時限メディアデータ・ビットストリームが得られるように、有利に順序づけられる。

【0071】

次に、ステップ604では、図3および図4について説明したように、サーバーは、ISO BMFF表現による時間的期間を含む初期化セグメントファイルおよびメディアセグメントファイルを生成し記憶する。全ての時限メディアデータトラック（例えばビデオトラック）、複合トラック、およびタイルトラックは、別々のメディアセグメントファイルに記憶される。

10

【0072】

それから、サーバーは、要求に応じて、クライアントデバイスに初期化セグメントファイルおよびメディアセグメントファイルを提供する（ステップ606）。サーバーは、HTTP要求に応答する通常のHTTPサーバーであってもよい。

【0073】

HTTPストリーミングのコンテキストにおいては、そして、好適な実施形態においては、サーバーから入手可能なメディアプレゼンテーションを記述するマニフェストファイルにクライアントデバイスがアクセスすることを仮定している。このマニフェストファイルは、クライアントデバイスが、サーバーから最初に初期化セグメントを要求し、それからメディアセグメントファイルを要求することによって、メディアプレゼンテーションをストリーミングするのに十分な情報（メディア特性およびセグメントのリスト）を提供する。

20

【0074】

クライアントデバイス側でROIを選択すると、通常はポインティングデバイスなどの選択手段を有するディスプレイ上で、タイル化されたビデオをストリーミングする間に、選択したROIに対応するタイルが決定される（ステップ608）。

【0075】

次に、時間的期間ごとに、スケーラブルなメディアデータの場合には、クライアントデバイスは、従属する層（ステップ610）に対応するセグメントファイルをダウンロードする要求をサーバーに送信する。特定の実施形態によれば、従属する層は、その層が従属する層の前にダウンロードされる。例えば、ベース層セグメントファイルは、拡張層セグメントファイルの前にダウンロードされる。

30

【0076】

続くステップにおいて、クライアントデバイスは、選択されたタイルに対応するメディアセグメントファイルをダウンロードする要求をサーバーに送信し（ステップ612）、複合トラックに対応するメディアセグメントファイルをダウンロードする要求をサーバーに送信する（ステップ614）。

【0077】

次に、ダウンロードされたセグメントファイルは、選択されたROIに対応し、ISO BMFF標準規格に合致する、有効な（復号化可能な）時限メディアデータ・ビットストリームを構築するために、クライアントデバイスによって連結される（ステップ616）。

40

【0078】

ステップ616については、図7を参照して詳細に説明する。

【0079】

図7は、クライアントデバイスによって受け取られた連結されたメディアセグメントからの有効な時限メディアデータ・ビットストリームの生成、すなわち、複合トラックがクライアントデバイスによって再生される場合に、選択されたタイルおよび1つの複合トラックから復号化可能なビットストリームの生成を示すフローチャートである。

【0080】

50

第1のステップ(ステップ700)において、クライアントデバイスは以前要求されたメディアセグメントファイルを受け取り(例えば図6のステップ612、614、および616)、少なくとも1つのメディアセグメントファイルが受け取られたか否かを決定するためにテストが実行される(ステップ702)。メディアセグメントファイルが受け取られなかった場合には、処理が終了する。

【0081】

少なくとも1つのメディアセグメントファイルが受け取られた場合で、かつ、受け取ったメディアセグメントファイルが複合トラックを含まない(すなわち、それらが従属するトラック、例えばより低いレベルの層またはタイルトラックを含む)場合には、それらは後で使用するためにバッファされる。

10

【0082】

逆に、少なくとも1つのメディアセグメントが受け取られた場合で、かつ、受け取ったメディアセグメントファイルが複合トラックを含む場合には、

複合トラックに対応するビットストリーム(通常はNALユニット)は、受け取ったメディアセグメントのメディアデータボックス「mdat」から、第1のデータの項目(または、受け取ったメディアセグメントの少なくとも1つのデータの項目が処理された場合には、次のデータの項目、通常はNALユニット)を抽出するために解析される(ステップ704)。解析するさらなるデータの項目がメディアデータボックスにない場合には、新しいメディアセグメントファイルを受け取るために、処理はステップ700に戻る(ステップ706)。

20

【0083】

次に、抽出されたデータの項目(例えば、抽出されたNALユニット)がエクストラクタに対応するか否かを決定するために、テストが実行される(ステップ708)。抽出されたデータの項目がエクストラクタに対応しない場合には、それはビデオデコーダによってさらに復号化するように戻される(ステップ710)。逆に、抽出されたデータの項目がエクストラクタである場合には、それが参照しているデータの項目と置き換えなければならない。そのために、エクストラクタのパラメータの値は、その構造から取得される(ステップ712)。上述したように、エクストラクタは、別のトラックからデータを抽出するために必要な全てのパラメータ値を含む(例えば、トラック参照インデックス、サンプルオフセット、データオフセット、およびデータ長)。

30

【0084】

一旦参照されたトラックの識別子が識別されると、参照されたトラックがステップ700でバッファされた1組のメディアセグメントファイルで利用可能か否かを決定するために、テストが実行される(ステップ714)。クライアントデバイスが選択された関心領域に対応するメディアセグメントファイルだけをダウンロードするので、いくつかのタイルトラックは失われていることを想起すべきである。

【0085】

参照されたトラックがステップ700でバッファされた1組のメディアセグメントファイルで利用可能な場合には、エクストラクタはそれが参照しているデータと置き換えられ(ステップ716)、ビットストリームは復号化するためにビデオデコーダに送信される(ステップ710)。

40

【0086】

参照されたトラックがステップ700でバッファされた1組のメディアセグメントファイルで利用できない場合には、エクストラクタで参照されるデータの欠如はISO BMF標準規格により致命的エラーとなるので、特定のステップを実行しなければならない。参照されたトラックがタイルトラックであるか否か(参照されたトラックは従属するスケラビリティ層に対応することができる)、およびエクストラクタがタイルタイプであるか否かを決定するために、テストが実行される(ステップ718)。

【0087】

参照されたトラックがタイルトラックでない場合、または、エクストラクタがタイルタ

50

イブでない場合には、標準致命的エラーが検出される。逆に、参照されたトラックがタイルトラックである場合で、かつ、エクストラクタがタイルタイプである場合には、エクストラクタは除去されるか（ステップ 722）、あるいは、エクストラクタは、時限メディアデータ・ビットストリームを符号化するために用いる符号化フォーマットに応じて（ステップ 720）、失われたタイルの「スキップされた」データを含む代替りの「パディングトラック」または「パディングボックス」からのパディングと置き換えられる（ステップ 724）。ここで、「スキップされた」データは、同じスケラブルな層に属するか、または別のスケラブルな層に属する、前に復号化された画像から取得された他の画素データと置き換えられている現行の画像の失われた画素データを表す。スキップされたデータは、一般に少なくとも 1 つのフラグによって表される。例えば、H E V C ビデオ圧縮形式を考慮すると、パディングデータは 1 にセットされたスキップフラグによって符号化された符号化ユニットを排他的に含む 1 つまたは複数の N A L U であってもよい。

10

【0088】

次に、ビットストリームは、復号化するためにビデオデコーダに送信されて（ステップ 710）、表示され、そして、続くデータの項目を扱うために、処理はステップ 704 に戻る。

【0089】

図 8 は、1 つまたは複数の実施形態のステップを実行することができる、サーバーまたはクライアントデバイス 800 のブロック図を示す。

【0090】

好ましくは、デバイス 800 は、通信バス 802、デバイスの電源投入時にプログラム R O M 806 からの命令、および電源投入後のメインメモリ 808 からのソフトウェアアプリケーションに関する命令を実行することができる中央処理ユニット（C P U）804 を含む。メインメモリ 808 は、例えば、通信バス 802 を介して C P U 804 の作業領域として機能するランダムアクセスメモリ（R A M）タイプであり、その記憶容量は拡張ポート（図示せず）に接続される任意の R A M によって拡張することができる。例えば、ソフトウェアアプリケーションに関する命令は、ハードディスク（H D）810 またはプログラム R O M 806 からメインメモリ 808 にロードすることができる。このようなソフトウェアアプリケーションが C P U 804 によって実行されると、図 6 a に関して説明したステップがサーバーで実行され、図 6 b および図 7 に関して説明したステップがクライアントデバイスで実行される。

20

30

【0091】

符号 812 は、デバイス 800 の通信ネットワーク 814 への接続を可能にするネットワークインターフェースである。ソフトウェアアプリケーションは、C P U 804 によって実行されると、ネットワークインターフェースを通して受け取った要求に反応して、他のデバイスに対してネットワークを介してデータストリームおよび要求を提供するように適合する。

【0092】

符号 816 は、ユーザーに情報を表示し、および / もしくはユーザーから入力を受け取るためのユーザーインターフェースを表す。

40

【0093】

変形例として、マルチメディア・ビットストリームの受信または送信を管理するためのデバイス 800 は、図 6 a、図 6 b、および図 7 に関して説明したような方法を実施することができる 1 つまたは複数の専用の集積回路（A S I C）から構成することができることを、ここで指摘すべきである。これらの集積回路は、例示的なものであって、非限定的に、ビデオシーケンスを生成もしくは表示するための、および / もしくは音声シーケンス聞くための装置に集積化される。

【0094】

上述したように、本発明の実施形態は、特に、H E V C として知られているビデオフォーマットに適用することができる。

50

【 0 0 9 5 】

H E V C 標準規格によれば、画像は、タイル、スライス、およびスライスセグメントに空間的に分割することができる。この標準規格では、タイルは、水平境界および垂直境界（すなわち、行および列）によって規定される画像の長方形領域に対応する。それは、整数個の符号化ツリーユニット（C T U）を含む。したがって、例えば、関心領域の位置およびサイズを定義することによって関心領域を識別するために、タイルを効率的に用いることができる。しかしながら、H E V C ビットストリームの構造、ならびにネットワークアブストラクト層（N A L）ユニットとしてのそのカプセル化は、タイルを考慮して構成されておらず、スライスに基づいている。

【 0 0 9 6 】

H E V C 標準規格では、スライスはスライスセグメントの組であって、1組のスライスセグメントの最初のスライスセグメントは、独立なスライスセグメント、すなわち、ヘッダー内に記憶された一般情報が別のスライスセグメントの一般情報を参照しないスライスセグメントである。1組のスライスセグメントの他のスライスセグメントは、存在する場合、従属するスライスセグメント（すなわち、ヘッダー内に記憶された一般情報が独立なスライスセグメントの一般情報を参照するスライスセグメント）である。

【 0 0 9 7 】

スライスセグメントは、整数個の連続的な（ラスタ走査順の）符号化ツリーユニットを含む。したがって、スライスセグメントは長方形の形状であっても、そうでなくてもよく、したがって関心領域を表すことには適していない。それは、スライスセグメントデータが後に続くスライスセグメントヘッダーの形式の下で、H E V C ビットストリームに符号化される。独立なスライスセグメントおよび従属するスライスセグメントは、それらのヘッダーによって区別される。すなわち、従属するスライスセグメントは独立なスライスセグメントに依存するので、そのヘッダーの情報量は独立なスライスセグメントのものより小さい。独立なスライスセグメントおよび従属するスライスセグメントは、タイルを定義するために用いられる対応するビットストリームのエントリ点、またはエントロピー復号化同期点としてのエントリ点のリストを含む。

【 0 0 9 8 】

図 9 は、図 9 a、図 9 b、および図 9 c を含み、タイルおよびスライスセグメントの例を示す。より正確には、図 9 a は、垂直境界 9 0 5 - 1 および 9 0 5 - 2 ならびに水平境界 9 1 0 - 1 および 9 1 0 - 2 によって 9 つの部分に分割された画像（9 0 0）を示す。符号 9 1 5 - 1 ~ 9 1 5 - 9 で示す 9 つの部分の各々は、特定のタイルを表す。

【 0 0 9 9 】

図 9 b は、垂直境界 9 0 5 ' によって区切られた 2 つの垂直タイルを含む画像（9 0 0 '）を示す。画像 9 0 0 ' は、5 つのスライスセグメント、すなわち 1 つの独立なスライスセグメント 9 2 0 - 1（ハッチング線で表す）および 4 つの従属するスライスセグメント 9 2 0 - 2 ~ 9 2 0 - 5 を含む単一のスライス（符号なし）を含む。

【 0 1 0 0 】

図 9 c は、垂直境界 9 0 5 ' ' によって区切られた 2 つの垂直タイルを含む画像（9 0 0 ' '）を示す。左のタイルは、2 つのスライス、すなわち、1 つの独立なスライスセグメント（9 2 0 ' - 1）および 1 つの従属するスライスセグメント（9 2 0 ' - 2）を含む第 1 のスライスと、1 つの独立なスライスセグメント（9 2 0 ' - 3）および 1 つの従属するスライスセグメント（9 2 0 ' - 4）を含む第 2 のスライスと、を含む。右のタイルは、1 つの独立なスライスセグメント（9 2 0 ' - 5）と、1 つの従属するスライスセグメント（9 2 0 ' - 6）と、を含む 1 つのスライスを含む。

【 0 1 0 1 】

H E V C 標準規格によれば、スライスセグメントは、以下のように要約することができるルールに従って、タイルにリンクされる（一方または両方の条件が満足されなければならない）。

【 0 1 0 2 】

1つのスライスセグメントの全てのCTUは、同じタイルに属する（すなわち、1つのスライスセグメントはいくつかのタイルに属することができない）。

【0103】

そして、1つのタイルの全てのCTUは、同じスライスセグメントに属する（すなわち、1つのタイルはいくつかのスライスセグメントに分割することができるが、ただし、これらのスライスセグメントの各々はそのタイルだけに属する）。

【0104】

明確にするため、以下では、1つのタイルがただ1つの独立なスライスセグメントを有する1つのスライスを含むと考える。しかし、本発明の実施形態は、図9bおよび図9cに示したように、他の構成によっても実施することができる。

10

【0105】

上述したように、タイルは関心領域に対する適切なサポートとみなすことができるが、スライスセグメントは通信ネットワーク上の輸送のためにNALユニットに実際に配置され、アクセスユニット（すなわちファイルフォーマットレベルの符号化画像またはサンプル）を形成するように集められた実体である。

【0106】

HEVC標準規格によれば、NALユニットのタイプが以下のように定義することができるNALユニットヘッダーの2バイトに符号化されることを想起すべきである。

```
nal_unit_header ( ) {
    forbidden_zero_bit
    nal_unit_type
    nuh_layer_id
    nuh_temporal_id_plus1
}
```

20

【0107】

スライスセグメントを符号化するために用いられるNALユニットは、スライスセグメント・アドレスシンタックス要素のおかげで、スライスセグメントの最初のCTUのアドレスを示すスライスセグメントヘッダーを含む。このようなスライスセグメントヘッダーは、以下のように定義することができる。

```
slice_segment_header ( ) {
    first_slice_segment_in_pic_flag
    if (nal_unit_type >= BLA_W_LP && nal_unit_type <= RSV_IRAP_VCL23)
        no_output_of_prior_pics_flag
    slice_pic_parameter_set_id
    if (!first_slice_segment_in_pic_flag) {
        if (dependent_slice_segments_enabled_flag)
            dependent_slice_segment_flag
            slice_segment_address
    }
    if (!dependent_slice_segment_flag) {
        [ ... ]
    }
}
```

30

40

【0108】

タイル化情報は、PPS（画像パラメータセット）NALユニットで提供される。それから、スライスセグメントとタイルとの間の関係は、これらのパラメータから推定することができる。

【0109】

空間的予測はタイル境界上でリセットされる（定義によって）が、タイルが参照フレー

50

ムの異なるタイルからの時間的予測を用いることは妨げられない。したがって、独立なタイルを構築するために、予測ユニットのための運動ベクトルは、符号化の間、参照フレームの同じ位置に配置されたタイルに留まるように、タイル内部に有利に拘束される。さらに、1つのタイルだけを復号化する場合に誤差変動がもたらされないように、ループ内フィルタ（非ブロッキングおよびサンプル適応可能なオフセット（SAO）フィルタ）は、好ましくはタイル境界上で非アクティブである。ループ内フィルタのこのような制御がHEVC標準規格において利用できる点に留意する必要がある。それは、`loop_filter_across_tiles_enabled_flag`として知られるフラグにより、スライスセグメントヘッダーにセットされる。このフラグをゼロに明示的にセットすることによって、タイル境界の画素は、隣接するタイルの境界にある画素に依存することができない。運動ベクトルおよびループ内フィルタに関するこれらの2つの条件が満たされる場合には、タイルは「独立に復号化可能なタイル」または「独立なタイル」とみなすことができる。

10

【0110】

ビデオビットストリームが1組の独立なタイルとして符号化されると、それは、参照データを失うリスクまたは再構築誤差の伝搬のリスクを伴わずに、1つのフレームから別のフレームまでタイルベースの復号化を可能にする。それから、この構成は、例えば、図2（タイル3および7を含む）に示した関心領域に対応することができる当初のビデオの空間部分だけを再構築することを可能にする。このような構成は、タイルベースの復号化は信頼性が高いことを示すために、ビデオビットストリームの補足情報として示すことができる。

20

【0111】

本発明の一実施形態によれば、HTTPストリーミングのコンテキストのタイルに対する効率的なアクセスは、HEVC標準規格に適用されるISO BMFFファイルフォーマットを用いて提供される。したがって、符号化される独立なタイルの各々（例えば図2に示した12のタイルの各々）は、図10に関して後述するように「タイルトラック」と呼ばれる特定のトラックによって表される。

【0112】

これらのタイルトラックは、図10に示すように、フルフレームのHEVCビットストリームに対応する複合トラックにおいて参照される（トラックの各々の定義を含むムービーボックス「moov」のトラック参照ボックス「tref」を介して）。各タイルトラックは、NALユニットにパッケージされた圧縮ビデオデータを含む。複合トラックは、初期化データに対応する様々なパラメータセット（例えばビデオパラメータセット、シーケンスパラメータセット、および/または画像パラメータセット）を含む。それは、特定のタイプのNALユニットであるエクストラクタも含む。

30

【0113】

上述したように、エクストラクタは、以下のシンタックスを有するファイルフォーマット内部構造であり得る。

```
class aligned(8) Extractor {
    NALUnitHeader();
    unsigned int(8) track_ref_index;
    signed int(8) sample_offset;
    unsigned int((lengthSizeMinusOne +
1) * 8)
        data_offset;
    unsigned int((lengthSizeMinusOne +
1) * 8)
        data_length;
}
```

40

【0114】

50

エクストラクタは、他のトラックからのデータへのポインタまたは参照として働き、両方のトラックでデータを複製する代わりに、従属するトラックを参照するコンパクトなトラックを構築することを可能にする。エクストラクタは、好ましくは、NALユニットシンタックスを用いる。したがって、それはNALユニットヘッダーと同一の構造を有するヘッダーを含み、特にNALユニットタイプに関係する情報を含む。このNALユニットタイプは、例えば、現在HEVCの予約されたNALユニットタイプに対応する値「47」にセットされる。ヘッダーの後には、エクストラクタによって参照されるトラックに対応するトラックの識別子(トラックid)を含むtreefボックスのエントリを読み出すことを可能にする、トラック参照ボックス(treef)のインデックス(トラック参照インデックスで示す)が続く。第3のパラメータは、現行のサンプルと比較してエクストラクタによって参照されるサンプルの時間オフセット(サンプルオフセット)である。第4および第5のパラメータ(データオフセットおよびデータ長で示す)は、コピーする位置(好ましくはバイト単位で)およびコピーするデータ量(値0は参照されたNALユニット全体のコピーを示すために確保される)をそれぞれ提供する。

【0115】

図10は、本発明の一実施形態による、複合トラックおよび独立タイルトラックを含む1組のトラックとしてHEVCビットストリームをカプセル化する例を示す。例示のために、カプセル化されたビットストリームは、図2に模式的に示したビデオシーケンスに対応する。

【0116】

図示するように、カプセル化されたビットストリーム1000は、複合トラック1015および12のタイルトラック1020-1~1020-12(タイルトラック1020-1~1020-12の各々は、ビデオシーケンスの1つのタイルに関係する)を表すトラックおよびメディアセグメントファイル1010の定義を提供する、ムービーボックス(「moov」)を含む初期化セグメントファイル1005を含む。

【0117】

複合トラック1015は、DASH標準規格の規定により、セグメントタイプボックス「styp」(図示せず)、トラックセグメントタイプおよび識別子などのメタデータを含む少なくとも1つのムービーフラグメントボックス「moof」1025、ならびに、ビデオデータサンプルごとに、PPSおよびビデオデータへの参照を含む少なくとも1つのメディアデータボックス「mdat」1030を含む。

【0118】

同様に、タイルトラック1020-1~1020-12の各々は、セグメントタイプボックス「styp」(図示せず)、トラックセグメントタイプおよび識別子などのメタデータを含む少なくとも1つのムービーフラグメントボックス「moof」、ならびに、NALユニット(NALU)にパッケージされた圧縮ビデオデータを含む少なくとも1つのメディアデータボックス「mdat」を含む。

【0119】

タイルトラック1020-1~1020-12は、識別子2~13を有し、初期化セグメントファイル1005の(より正確には、識別子id=1を有する複合トラックの定義において、初期化セグメントファイル1005のムービーボックス「moov」の)トラック参照ボックス「treef」1035で参照される。

【0120】

図示するように、複合トラック1015は、他のトラックからのデータへのポインタまたは参照として働くエクストラクタを含む。例示のために、いくつかのパラメータの中で、複合トラック1015のエクストラクタ1035-1および1035-pに対応する、タイルトラックのインデックス(トラック参照インデックス)、データのオフセット(データオフセット)、およびデータの長さ(データ長)を示す。

【0121】

さらに例示のために、複合トラック1015のNALユニット1035-1が処理され

10

20

30

40

50

ると、それがエクストラクタタイプのNALユニット(NALユニットヘッダーは16進数5E00に等しい)を表すと決定される。したがって、それは、対応する圧縮ビデオデータを回復するために処理される。そのために、そのタイルトラックインデックス(すなわちトラック参照インデックス=1)が取得される。このインデックスから、初期化セグメントファイル1005に記憶されたタイルトラック定義からタイルトラック識別子を回復することができる。所与の例では、インデックスが1に等しいので、treefボックスの最初のタイルトラック識別子が選択される(id=2)。次に、この識別子は、対応するタイルトラックにアクセスするために用いられ、それから、エクストラクタ1035-1のデータオフセット(すなわち、情報源として用いられる識別されたトラックのサンプルの相対インデックス)およびデータ長(すなわち、コピーするバイト数、例えばデータ長=0の場合はNALU全体)パラメータを用いて、圧縮ビデオデータがタイルトラック1020-1(すなわち、所与の例では符号化スライスセグメントNALU1040)から抽出される。

10

【0122】

処理された後に、エクストラクタは、それが参照するデータと置き換えられる。図10に示す例によれば、エクストラクタ1035-1の解析および処理は、符号化スライスセグメントNALU1040による置き換えをもたらし、このようにして、HEVCに準拠するビットストリームを形成する。

【0123】

HEVCエクストラクタのパラメータを記憶するために用いられるセマンティックが、SVC標準規格で定義されるものに近くなり得る点に留意する必要がある。したがって、HEVC NALユニットを参照するエクストラクタについて、以下のことが当てはまる。

20

【0124】

禁止ゼロビット(forbidden_zero_bit)として知られるパラメータは、ISO/IEC 23008-2の規定により設定される。

【0125】

nalユニットタイプ(nal_unit_type)として知られるパラメータは、47(現行のFDISの予約されたコード)に設定される。

【0126】

nuh層id(nuh_layer_id)およびnuh時間idプラス1(nuh_temporal_id_plus1)として知られるパラメータは、エクストラクタによって参照される第1のNALUからコピーされる(HEVC NALユニットを参照するHEVCトラックのエクストラクタは、異なるnuh層id(nuh_layer_id)およびnuh時間idプラス1(nuh_temporal_id_plus1)値を有するいくつかのNALユニットを参照しない)。

30

【0127】

サンプルオフセット(sample_offset)として知られるパラメータは、0に設定される。

【0128】

本発明の実施形態によれば、タイルトラックは、空間的に配置されたデータ(例えばHEVCタイル)のインデクセーションをサポートするために用いる特定のトラックとみなされる。したがって、特に、フルビデオにおけるタイルの位置(水平座標および垂直座標)およびタイルのサイズ(幅および高さ)を示し、タイルトラックがアーチファクトなしに復号化することができるか否かを示す情報を提供するための、特定の通知がトラックごとに必要となる。このような通知は、いくつかの実施形態により達成することができる。

40

【0129】

上述したものと同様な図11に示す特定の実施形態によれば、通知はトラックレベルで行うことができる。

【0130】

50

図 11 は、H E V C ビットストリームをカプセル化するための、第 1 の実施形態による、メディアデータトラックを用いる初期化セグメントファイルのブロック図の例を示す。それは、図 10 を参照して説明した例に基づいている。

【0131】

初期化セグメントファイル 1100 およびメディアセグメントファイル（図示せず）は、I S O ベースメディアファイルフォーマットの H E V C 標準規格に準拠するビデオビットストリームを再構成しカプセル化するために用いられる。トラックは、独立にストリーミングされる。

【0132】

上述したように、他のメディアセグメントファイルにカプセル化される時限メディアデータ・ビットストリームを定義するのに必要な全てのメタデータを送信するために、初期化セグメントファイルを用いる。図 11 に示すように、初期化セグメントファイル 1100 は、ファイルタイプボックス「f t y p」1105 およびムービーボックス「m o o v」1110 を含む。ファイルタイプボックス 1105 は、好ましくは、セグメントファイルがどの I S O B M F 仕様に対応するかを識別して、その仕様のバージョン番号を示す。ムービーボックス「m o o v」1110 は、メディアセグメントファイルに記憶されたプレゼンテーションを記述する全てのメタデータ、および特にプレゼンテーションで利用可能な全てのトラックを提供する。

【0133】

ムービーボックス 1110 は、所与の例では、1 つの複合トラック（1115 - 1）および 12 のタイルトラック（1115 - 2 ~ 1115 - 13）を含むトラック（「トラック」ボックス 1115 - 1 ~ 1115 - 13）の各々の定義を含む。

【0134】

各トラックボックスは、少なくとも、一般的に符号 1120 で示すトラックヘッダーボックス「t k h d」および一般的に符号 1125 で示すトラックメディアボックス「m d i a」を含む。トラックが他のトラックからのデータに依存する場合には、トラック参照ボックス「t r e f」もある。図示するように、識別子トラック I D = 1 を有する複合トラックは、トラックが識別子トラック I D = 2 ~ 13 を有するタイルトラックからのデータに依存することを示すトラック参照ボックス「t r e f」1130 を含む。

【0135】

上述したように、時限メディアデータ・ビットストリームをカプセル化するために用いる I S O B M F F 仕様に応じて、他のボックスが必須であってもよいし、あるいは任意であってもよい点に留意する必要がある。しかし、本発明の実施形態は、適用可能なこれらのボックスに依存しないので、それらについてはここで提示しない。

【0136】

図 11 を参照して説明する実施形態によれば、フルビデオにおけるタイルの位置、タイルのサイズ、および、タイルトラックがアーチファクトなしに復号化することができることの通知は、「m o o v」ボックス（1110）で、各トラック定義において、トラックヘッダーボックス「t k h d」（1120）およびメディア情報ボックス「m d i a」（1125）のボックスを用いて、カプセル化される H E V C ビットストリーム全体について一度行われる。

【0137】

タイルの位置は、タイルメディアハンドラエントリ（T i l e M e d i a H a n d l e r E n t r y）ボックスと呼ばれるメディアヘッダー情報ボックス 1135 の新しいタイプ、または水平オフセットおよび垂直オフセット（h o r i z o n t a l _ o f f s e t および v e r t i c a l _ o f f s e t）を定義する「t m h d」（1140）に配置される。

【0138】

考慮されたタイルのビジュアルプレゼンテーションのサイズは、既存のトラックヘッダーの幅および高さパラメータに入力される。必要であれば、考慮されたタイルの現実の画

10

20

30

40

50

素寸法は、サンプル記述に記録することができる（サンプル記述ボックス（Sample Description Box）として知られるボックスのビジュアルサンプルエントリ（Visual Sample Entry（））として知られるパラメータを介して）。

【0139】

最後に、特定のトラック（タイルトラック）を通知することは、ハンドラボックス「hdlr」（1145）のトラック（「tile」）の新しいタイプを定義することによって、明示的に行われる。

【0140】

初期化セグメントファイル1100のムービーボックス「moov」1110は、「mvex」ボックス1150をさらに含む。カプセル化されたファイルにアクセスしているクライアントに、ムービーフラグメントが存在するという情報を与えるために、このボックスが用いられる。それは、初期化セグメントファイルにおいて、プレゼンテーションで最も長いトラックの継続時間を指定することを可能にする。それはプレゼンテーション継続時間の計算をより簡素にし、各ムービーフラグメント継続時間の検査を回避する。図示するように、mvexボックス1150は、各トラック（すなわちタイルトラックおよび複合トラック）の全てのフラグメントに共通な情報、例えばトラック識別子およびトラックのサンプルのデフォルトサイズの重複を回避するために、トラックごとにトラック拡張ボックスを含む。

【0141】

タイルトラックを通知するこのような実施形態が、通常のビデオトラックの通知と比較して著しいオーバーヘッドをもたらさない点に留意する必要がある。さらに、それは、プレゼンテーションが分割されるムービーフラグメントの数から独立した状態を維持する。しかし、それは、ビデオシーケンスに沿ってタイル構成を変化させることには適応しない。

【0142】

ビデオシーケンスに沿ってタイル構成の変化を扱うのに適する別の特定の実施形態によれば、ISO BMFF標準規格のサンプルグルーピング機構を用いてサンプルレベルで通知が行われる。

【0143】

このようなサンプルグルーピング機構は、トラックのサンプルの分割を表すために用いられる。それらは、2つのボックス、すなわち、サンプルのサンプルグループへの割当てを記述するグループへのサンプル（Sample To Group）ボックス（「sbgp」）、および、特定のサンプルグループ内のサンプルの共通の特性を記述するサンプルグループ記述（Sample Group Description）ボックス（「sgpd」）の使用に依存している。サンプルグルーピングの特定のタイプは、タイプフィールド（「グルーピングタイプ」）を介して、1つのグループへのサンプル（Sample To Group）ボックスおよび1つのサンプルグループ記述（Sample Group Description）ボックスの組み合わせによって定義される。複数のサンプルグルーピングインスタンス（すなわちSample To GroupおよびSample Group Descriptionボックスのペア）が、種々のグルーピング基準に基づいて存在することができる。

【0144】

本発明の実施形態によれば、サンプルのタイル化に関する新しいグルーピング基準が定義される。「tile」と呼ばれる、この新しいグルーピングタイプは、タイルの特性を記述し、標準ビジュアルサンプルグループエントリ（Visual Sample Group Entry）から導出される。それは、タイル領域サンプルグループエントリ（Tile Region Sample Group Entry）またはHEVC空間エントリ（HEVC Spatial Entry）と呼ぶことができ、以下のように定義される。

```
class HEVC SpatialEntry() extends VisualS
```

```

ampleGroupEntry ( ' t r s g ' ) {
unsigned int ( 3 2 )   t i l e I D ;
unsigned int ( 1 6 )   h o r i z o n t a l _ o f f s e t ;
        unsigned int ( 1 6 )   v e r t i c a l _ o f f s e t ;
        unsigned int ( 1 6 )   r e g i o n _ w i d t h ;
        unsigned int ( 1 6 )   r e g i o n _ h e i g h t ;
        unsigned int ( 2 )     i n d e p e n d e n t ;
        unsigned int ( 6 )     r e s e r v e d = 0 ;
}

```

【 0 1 4 5 】

10

グループエントリのこの新しいタイプによれば、タイルID (t i l e I D) パラメータは、グループにより記述されたタイルの一意的な識別子である。水平オフセット (h o r i z o n t a l _ o f f s e t) および垂直オフセット (v e r t i c a l _ o f f s e t) パラメータは、ベース領域の輝度サンプルにおいて、HEVCフレームの左上画素に対して、タイルによって表される長方形領域の左上画素の水平オフセットおよび垂直オフセットを設定するためにそれぞれ用いられる。領域幅 (r e g i o n _ w i d t h) および領域高さ (r e g i o n _ h e i g h t) パラメータは、HEVCフレームの輝度サンプルにおいて、それぞれ、タイルによって表される長方形領域の幅および高さを設定するためにそれぞれ用いられる。独立 (i n d e p e n d e n t) パラメータは、独立なタイルの定義について上述したように、タイルが同じタイルだけに属するサンプルに関する復号化従属性を含むことを指定する2ビットワードである。例示のため、およびタイル構成を説明するためのSEIメッセージの標準的使用に関して、タイルセクション正確一致フラグ (t i l e _ s e c t i o n _ e x a c t _ m a t c h _ f l a g) として知られるフラグは、独立フラグの値をセットするために用いることができる。後者の意味は、以下のように設定することができる。

20

【 0 1 4 6 】

独立パラメータが0に等しい場合には、このタイルと、同じフレームまたは前のフレームの他のタイルとの間の符号化従属性は不明である。

【 0 1 4 7 】

独立パラメータが1に等しい場合には、このタイルと同じフレームの他のタイルとの間には空間的な符号化従属性はないが、このタイルと前のフレームの同じタイルIDを有するタイルとの間には符号化従属性があり得る。

30

【 0 1 4 8 】

独立パラメータが2に等しい場合には、このタイルと、同じフレームまたは前のフレームの同じタイルIDを有する他のタイルとの間には符号化従属性がない。

【 0 1 4 9 】

独立パラメータ値3は保留されている。

【 0 1 5 0 】

選択的に、タイルごとの平均ビットレートを記述するパラメータを、バンド幅に基づく適合のためにストリーミングクライアントに提供するように、タイル記述子に設定することができる。

40

【 0 1 5 1 】

この実施形態によれば、各タイルの特性は、タイルトラックごとに、「タイル」グループピングタイプおよびHEVC空間エントリ (H E V C S p a t i a l E n t r y) を有する1つのサンプルグループ記述 (S a m p l e G r o u p D e s c r i p t i o n) ボックス (「 s g p d 」) を定義することによって、ムービーヘッダー (「 m o o v 」 ボックス) に一度与えられる。それから、ISO BMFF標準規格によれば、サンプル数が前もって知られていないので、タイルトラックフラグメントの各サンプルをその特性と関係づけるために、グループへのサンプル (S a m p l e T o G r o u p) ボックスが各タイルトラックフラグメントで定義される。

50

【0152】

タイルのグリッドが時間とともに変化する場合には、新規なH E V C空間エントリ (H E V C S p a t i a l E n t r y) を有する新しいサンプルグループ記述 (S a m p l e G r o u p D e s c r i p t i o n) ボックス (「s g p d」) が、トラックフラグメントボックス (「t r a f」) で定義されて、グループへのサンプル (S a m p l e T o G r o u p) ボックス (「s b g p」) によって参照され得る。したがって、グリッドが時間とともに変化しない場合には、少なくとも1つのグループへのサンプル (S a m p l e T o G r o u p) ボックスは、タイルトラックおよびタイルトラックフラグメントごとに定義される。このボックスは、記述に関して、少なくとも28バイトを表す。2秒の継続時間のフラグメントを有する16のタイルを仮定すると、これは、グループへのサンプル (S a m p l e T o G r o u p) ボックスについてだけ、時間とともにタイル化構成を通知するための1792ビット/秒を表す。グリッドが時間とともに変化する場合には、コスト (データ量に関して) がより高くなる。後述するように、この補足的な初期化データ量は低減することができる。

10

【0153】

サンプルグループボックスが、個々のサンプルにメタデータを追加するために、効率的で拡張可能なツールを提供する点に留意する必要がある。しかし、所与のメタデータが所与のトラックの全てのサンプルについて有効であることは、これが各A Uの固定N A L Uパターンを有する (すなわち時間に沿った静的タイル化構成を有するタイル化記述子の場合であり得るので、非常に一般的なことである。

20

【0154】

これは、したがって、グループへのサンプル (S a m p l e T o G r o u p) ボックスを用いて、グループの各サンプルにフラグを立てることによって示すことができる。

【0155】

あるいは、タイルトラックごとの初期化データ量を低減するために、いくつかのサンプルグループは、「デフォルト」 (すなわち全てのサンプルに適用) のマークを付けることができる。このデフォルトグルーピングは、静的タイル化構成では、それが「m o o v」ボックスレベルで、シーケンス全体についてトラックごとに1回定義されるので、バイトに関して記述コストを制限する。

【0156】

そのために、グループ記述タイプの第2のバージョン (バージョン = 2) は、サンプルグループ記述ボックス (S a m p l e G r o u p D e s c r i p t i o n B o x) として知られるボックス (t r a f / s t b l ボックスごとに複数の S a m p l e G r o u p D e s c r i p t i o n B o x があってもよい) で用いることができ、参照されたサンプルグループが現行のトラックまたは現行のトラックフラグメントの全てのサンプルに当てはまることを示す (グルーピングタイプとして知られるパラメータを介して)。

30

【0157】

サンプルグループ記述ボックスの新しいバージョンは、以下のように定義することができる。

```
aligned(8) class SampleGroupDescriptionBox(
    unsigned int(32) handler_type) extends FullBox(
    'sgpd', version, 0) {
    unsigned int(32) grouping_type;
    if (version == 1) || (version == 2) {
        unsigned int(32) default_length;
    }
    unsigned int(32) entry_count;
    int i;
    for (i = 1; i <= entry_count; i++) {
        if (version != 0) {
            if (default_length == 0) {
```

40

50

```

        unsigned int(32) description_length
    th;
}
}
switch(handler_type) {
    case 'vide': // for video tracks
VisualSampleGroupEntry(grouping_type);
break;
    case 'soun': // for audio tracks
AudioSampleGroupEntry(grouping_type);
break;
    case 'hint': // for hint tracks
HintSampleGroupEntry(grouping_type);
break;
}
}

```

10

【0158】

サンプルグループ記述(SampleGroupDescription)ボックスのこの拡張タイプによれば、バージョンが2に等しい場合には、このグループ記述および任意のサンプルグループエントリ(SampleGroupEntry)は、現行のトラックまたはトラックフラグメントの全てのサンプルに当てはまる。グルーピングタイプ(grouping_type)パラメータは、このサンプルグループ記述と関係するグループへのサンプル(SampleToGroup)ボックスを識別する整数である。トラックまたはトラックフラグメントが、その初期化データにバージョン=2のサンプルグループ記述(SampleGroupDescription)ボックスを含む場合には、対応するグルーピングタイプを記述するグループへのサンプル(SampleToGroup)ボックスを、これらの初期化データに入れる必要がない(定義によって記述が全てのサンプルに適用されるので)という点に留意する必要がある。

20

【0159】

サンプルグループ記述ボックスの新しいバージョンのこの定義では、エントリ数(entry_count)は、続くテーブルのエントリ数を与える整数であり、デフォルト長(default_length)は、あらゆるグループエントリの長さ(長さが一定である場合)、またはそれが可変である場合にはゼロを示し、記述長(description_length)は、エントリごとにエントリの長さが変化し、したがってデフォルト長(default_length)がゼロである場合に、個々のグループエントリの長さを示す。

30

【0160】

この実施形態によれば、グルーピングタイプは、空間的ノタイルグルーピングを示すために、特定の値を有することができ、それは、例えば、「タイル」(0x74696C65)のASCIIコードに対応する16進数とすることができる。グルーピングタイプの同じ値を有するこのボックスの高々1つの発生が、1つのトラックについて存在する。

40

【0161】

時間とともに移動する適応可能なグリッドの場合には、グループへのサンプルボックスは同じままであって(すなわち「タイル」グルーピングタイプ)、全てのサンプルに適用し続ける。このように、サンプルグループ記述ボックスだけが、タイル化構成がmoov/trak/mdia/minf/stblボックスで通知されたデフォルト構成から変化したタイルトラックのトラックフラグメントにおいて更新されることを必要とする。これは、適応可能なタイルの通知コストを減らす。

【0162】

あるいは、さらに、タイルトラックごとの初期化データ量を減らすために(各タイルト

50

ラックフラグメントの `SampleToGroup` ボックスを繰り返すことを回避するために)、`'dsgrp'` で示す新しいデフォルトのグループへのサンプル (`DefaultSampleToGroups`) ボックス (または、その名称がどうであれ、同じセマンティックを有する別の類似のボックス) が、初期化情報の一部として各 `moov/trak` ボックスからのサンプルテーブル (`SampleTable`) ボックス (`'stbl'`) だけに含まれるように定義される。この新しいボックスは、トラック内の全てのサンプルに適用する 1 組のサンプルグループ記述を全てのサンプルに関係づける。

【0163】

新しいデフォルトのグループへのサンプル (`DefaultSampleToGroups`) ボックスは、以下のように定義することができる。

```
aligned(8) class DefaultSampleToGroups extends FullBox('dsgrp', version, 0) {
    unsigned int(32) entry_count;
    for (i = 1; i <= entry_count; i++) {
        unsigned int(32) grouping_type;
        if (version == 1) {
            unsigned int(32) grouping_type_parameter;
        }
        unsigned int(32) group_description_index;
    }
}
```

【0164】

ここで、エントリ数 (`entry_count`) パラメータは各サンプルに係するグループのリストのエントリ数を与え、グルーピングタイプ (`grouping_type`) パラメータは、グルーピングのタイプの識別子であって、サンプルグループ記述 (`SampleGroupDescription`) ボックスに記述される。例えば、特定の実施形態では、グルーピングタイプは、空間的 / タイルグルーピングを示す特定の値を有することができる。それは、例えば、`'タイル' (0x74696C65)` の `ASCII` コードに対応する 16 進数とすることができる。グループ記述インデックス (`group_description_index`) パラメータは、このグループのサンプルを記述するサンプルグループエントリのインデックスを与える整数である。このインデックスは、1 からサンプルグループ記述 (`SampleGroupDescription`) ボックスのサンプルグループエントリ数までの範囲に及び、あるいは、このサンプルがこのタイプの任意のグループのメンバーでないことを示すために値 0 をとる。最後に、グルーピングタイプパラメータ (`grouping_type_parameter`) パラメータは、グルーピングのサブタイプの指標である (グルーピングタイプによって用いられる場合)。

【0165】

これによって、タイルグルーピングが用いられてさえすれば (エントリ数 = 1)、ムービーフラグメントの数が何であっても、タイルごとに多くとも 32 バイトを用いて、トラックからの全てのサンプルが所与のグルーピングタイプの同じグループ記述に続くことを通知することが可能になる。時間とともに移動する適応可能なグリッドの場合には、新しいデフォルトのグループへのサンプル (`DefaultSampleToGroups`) ボックスおよび新しいサンプルグループ記述 (`SampleGroupDescription`) ボックスは、トラックフラグメントで定義することができる。新しいデフォルトのグループへのサンプル (`DefaultSampleToGroups`) ボックスは、以前の定義を置き換えて、新しいサンプルグループ記述 (`SampleGroupDescription`) ボックスの新しいタイル記述を参照する。したがって、グループへの

サンプル (Sample To Group) ボックスは、タイルグリッド定義が変化する場合を除いて、トラックフラグメントごとに定義されない。

【 0 1 6 6 】

さらに、ビデオシーケンスに沿ってタイル化構成の変化を扱うのに適する特定の実施形態によれば、新しいサンプルマップボックスを用いてサブサンプルレベルで通知が行われる。

【 0 1 6 7 】

タイル化構成に関するサンプル記述についてのこの実施形態は、単一のトラックのカプセル化されたタイル化 H E V C ビットストリームに適用される。それは、複合トラックのエクストラクタの解析の後で、M P E G - 4 標準規格に準拠する単一のビデオトラックにダウンロードされた空間部分を保存するために、タイル化 H E V C ビットストリームの M P E G - 4 カプセル化のための、またはクライアント側のアプリケーションを見いだすことができる。別のアプリケーションが、M P E G - 4 に準拠する標準的な単一トラックへの複合トラックの変換に向けられる場合には、エクストラクタをサポートしない M P E G - 4 パーサーを対象にする。

【 0 1 6 8 】

図 1 2 は、図 1 2 a および図 1 2 b を含み、フルビデオにおけるタイルの位置、タイルのサイズ、および様々なタイル化構成を扱うのに適したタイルトラックがサブサンプルレベルでアーチファクトなしに復号化できることの通知を示す。

【 0 1 6 9 】

図 1 2 a は、クライアントデバイス (例えばビデオプレーヤ) によって実行されるステップを示す。第 1 のステップ (ステップ 1 2 0 0) では、クライアントデバイスが初期化データをダウンロードし、または、ファイルがローカルファイルである場合には初期化データ、例えば M P E G - 4 標準規格に準拠するカプセル化されたビットストリームの初期化データ、通常は m o o v ボックスのコンテンツを読み込む。

【 0 1 7 0 】

これらの初期化データから、クライアントデバイスは、タイル情報が符号化されているトラックヘッダー情報を解析することができる (ステップ 1 2 0 5)。このタイル情報を用いて、ユーザーは、1 つまたは複数のタイルに対応することができるクライアントデバイスのグラフィックインターフェースを介して、関心領域を選択することができる (ステップ 1 2 1 0)。

【 0 1 7 1 】

対応するタイルトラックならびに複合トラックは、クライアントデバイスによって、ダウンロードされ、または読み込まれる (ステップ 1 2 1 5 および 1 2 2 0)。次に、複合トラックのエクストラクタは、単一のビデオトラックを取得するために、タイルトラックを用いて解析される (ステップ 1 2 2 5)。最後に、クライアントデバイスは、例えばサンプルテーブルボックス (Sample Table Box) において、タイル化記述を取得したビデオトラックに構築し追加する (ステップ 1 2 3 0)。

【 0 1 7 2 】

タイル化記述の例を図 1 2 b に示す。図示するように、タイル化記述 1 2 5 0 は、ムービーボックス「m o o f」1 2 5 5 およびデータボックス「m d a t」1 2 6 0 を含む。「m o o f」ボックス 1 2 5 5 は、種々のサンプルグループを記述するグループへのサンプル (Sample To Group) ボックス 1 2 6 5 を含む、トラックごとに 1 つのサンプルテーブル (Sample Table) ボックスと、各サンプルの N A L ユニット間のマッピングを記述するサンプルグループ記述ボックス 1 2 7 0 と、タイル記述を含むタイルおよびサンプルグループ記述ボックス 1 2 7 5 と、を含む。グループへのサンプルボックス 1 2 6 5 は、グループエントリであるタイルサンプルマップエントリ (Tile Sample Map Entry) の「t s g m」グループタイプを示す。

【 0 1 7 3 】

タイル N A L U マップエントリ (Tile N A L U Map Entry) グループエント

10

20

30

40

50

リ1270は、サンプルのNALユニットとタイルとの間のマッピングを定義する（これはこのような実施形態がサブサンプルレベルの通知を参照する理由である）。グルーピングタイプ（`grouping_type`）パラメータが「`tsgm`」に等しいこのボックスは、サンプルごとのNALユニット数を含む。

【0174】

タイルNALUマップエントリ（`TileNALUMapEntry`）ボックスは、以下のように定義することができる（図12bに示すように）。

```
class TileNALUMapEntry() extends VisualSampleGroupEntry('tsgm') {
  unsigned int(8) reserved = 0;
  unsigned int(8) entry_count;
  for (i = 1; i <= entry_count; i++)
    unsigned int(32) tileID;
}
```

10

【0175】

ここで、エントリ数（`entry_count`）はトラックサンプルのNALUの数を示し、タイルID（`tileID`）は、現行のトラックによって記述される空間タイルの一意的な識別子を提供する。

【0176】

20

タイルNALUマップエントリ（`TileNALUMapEntry`）ボックスは、サイズに関して以下のように最適化することもできる。

```
class TileNALUMapEntry() extends VisualSampleGroupEntry('tsgm') {
  unsigned int(6) reserved = 0;
  unsigned int(1) large_size;
  unsigned int(1) mode;
  if (large_size) {
    unsigned int(16) entry_count;
  } else {
    unsigned int(8) entry_count;
  }
  for (i = 1; i <= entry_count; i++)
    if (mode) {
      if (large_size) {
        unsigned int(16) NALU_start_number;
      } else {
        unsigned int(8) NALU_start_number;
      }
    }
  unsigned int(16) tileID;
}
```

30

40

【0177】

ここで、ラージサイズ（`large_size`）パラメータは、トラックサンプルのNALユニットエントリ数が8ビットで表現されるか、あるいは16ビットで表現されるかを示し、モード（`mode`）パラメータは、各NALユニットエントリが記述されているか（モードが設定されていない場合）、あるいはタイルIDに対応するNALユニットエントリだけが変化するか（モードが設定された場合）を示す。後者の場合には、NALユニット番号は、ラージサイズパラメータ（`large_size`）の値に応じて、16ビ

50

ットまたは8ビットに符号化される。

【0178】

それは、これらのNALユニットの各々と特定のタイルとの間のマッピングを指定する。上述したように、タイル記述は、サンプルグループ記述ボックス1275で提供され、各タイルが次々と記述される。

【0179】

所与の例は、1つのNALユニットがシーケンス継続時間に沿って1つのタイルおよびこのデータを含む、特定の場合である点に留意する必要がある。タイルデータがいくつかのNALユニットにわたって分割される場合には、タイルに対応する1組のNALユニットを記述するために、いくつかのアグリゲータが用いられる。NALユニットの数が時間とともに変化する場合には、いくつかのタイルサンプルマップ(tileSampleMap)エントリを定義することができて、グループへのサンプルは、代わりに、1つのフラグメントから別のフラグメントへ、トラックフラグメントヘッダーにおいて、グルーピングタイプを介して適切なタイルマップエントリを参照する。

【0180】

図10を参照して説明するデータカプセル化方式では、HEVCビットストリームは、実際に圧縮ビデオデータを含むタイルトラック1020-1~1020-2を示す複合トラック1015としてカプセル化される。複合トラックは、NALユニットが設定された異なるHEVCパラメータ(図10にPSで示す)に由来するコンフィギュレーションデータを含む。複合トラックの他の要素は、主にエクストラクタ(サンプルごとに1つ、およびタイルトラックごとに1つ)のリストであって、タイルトラックにカプセル化された圧縮ビデオデータを示す(初期化セグメントファイル1005のmovボックスに含まれるトラック参照ボックス('tref'))を介して)。

【0181】

ISO BMFF標準規格(標準規格のパート15)の現行の従属性の通知手段は、初期化セグメントファイル1005のmovボックスのトラックボックスの一部であるトラック参照ボックス'tref'に位置する。'tref'ボックスは、プレゼンテーションの含んでいるトラックから別のトラックへの参照を提供する。含んでいるトラックは、プレゼンテーションの複数の他のトラックを参照することができる。トラック間の従属性のタイプは、現行の標準規格では、2値すなわち'scal'または'sbas'を取ることができる参照タイプ(reference_type)パラメータによって指定される。'sbas'値はスケーラブルなベースを表す。それは、参照されたトラックがスケーラブルなプレゼンテーションの現行のトラックのスケーラブルなベーストラックであることを示す。'scal'値はスケーラビリティを表す。それは、スケーラブルな表現の種々の層を表すトラック間の関係を示す。それは、含んでいるトラックが参照されたトラックに依存することを意味する。

【0182】

図10を参照して説明する実施形態では、特定のスケーラビリティに関連する従属性はない。スケーラブルなビデオを考慮に入れることができる場合であっても、複合トラックとタイルトラックとの間の空間的従属性に、ここでは焦点が当てられる。これらの従属性は、例えば、複合トラック1015(id=1)に対応する初期化セグメントファイル1005のmovボックスのtrefボックスにおいて成されるような、新しい「タイル」値によって明示的に示すことができる。

【0183】

1つのトラックから他のトラックへの従属性の図示する例は、タイル(すなわち1つの複合トラックからタイルトラックへの従属性)を対象としているが、1つのトラックから1つまたはいくつかのトラックへの他のタイプの従属性も同様に扱うことができる。したがって、参照タイプ(reference_type)パラメータは、トラックが1つまたは複数のトラックに依存することを示すために、例えば、下位層従属性(例えば'sub1'値)を示すことができる。

【0184】

したがって、下位層トラックは、HEVC基本のビットストリームの一部を含むトラックとして定義することができ、他のHEVCNALユニットの復号化処理を害せずに廃棄することができる。上述したように、このような定義は、特に、スケーラブルなHEVCビットストリームの時間層、ならびにタイルトラックに当てはまる。下位層トラックに対応する各トラックは、所定の値にセットされた場合に、このHEVCトラックが下位層トラックであって、他のトラック、例えば複合トラックから参照されるNALユニットを含むだけである（すなわち、このHEVCトラックは表示可能でない）ことを示すビット（またはフラグ）を用いて、HEVC構成（HEVCConfiguration）レコード（すなわちSampleTableBox）にマークすることができる。このビットまたはフラグの値が反対の値を有する場合には、それは、このHEVCトラックが初期化データも含む下位層トラックである（すなわち、このHEVCトラックは表示可能である）ことを示す。例えば、現行のHEVCデコーダ構成レコード（HEVCDecoderConfigurationRecord）ボックスの予約されたビットを用いることができる。

10

【0185】

図12を参照して説明する特定の実施形態によれば、各タイルトラックは、標準ビデオトラックとして再生可能である。

【0186】

図13は、本発明の一実施形態による、標準ビデオトラックとして再生可能な複合トラックおよび独立タイルトラックを含む1組のトラックとしてHEVCビットストリームをカプセル化する例を示す。例示のために、カプセル化されたビットストリームは、図2に模式的に示したビデオシーケンスに対応する。

20

【0187】

図13に示すHEVCビットストリームのカプセル化は、主に、各タイルトラックが初期化およびコンフィギュレーションデータの回復を可能にする特定のエクストラクタを含むという点で、図10に示すものと異なっている。

【0188】

図示するように、タイルトラック1300-1~1300-12の各々は、初期化およびコンフィギュレーションデータを表す、複合トラック1310のNALユニット（PSで示す）がセットされたHEVCパラメータを示すエクストラクタ1305-1~1305-12を含み、HEVC標準規格によれば、これらの初期化およびコンフィギュレーションデータは、通常、HEVCビットストリームの様々なパラメータセットに対応することが想起される。したがって、このような初期化およびコンフィギュレーションデータは、各タイルトラックを通常のビデオトラックとして再生可能にする。

30

【0189】

各タイルトラックに追加されたエクストラクタは、各タイルトラックのメディアデータボックス「mdat」の始めに、すなわちビデオデータサンプルの前に有利に位置する。

【0190】

符号1315-1~1315-12で示す、タイルトラック（1300-1~1300-12）から複合トラック（1310）へのこれらの従属性は、例えば、タイルトラックと関係する「tref」ボックス1320-1~1320-12の参照タイプ（reference_type）パラメータにおいて（初期化セグメントファイル1325のムービーボックス「moov」において）、通知されなければならない。この実施形態によれば、パラメータセットを含むトラックは、HEVCベーストラック「hbas」とみなされる（これは、スケーラブルなプレゼンテーションにおいて最も低い動作点を含むトラックが「スケーラブルなベーストラック」「sbas」とみなされるSVCの場合に近い）。図示するように、ベーストラックに依存するトラック（すなわち、識別子id=1を有する複合トラック1310に依存する、識別子id=2~12を有するタイルトラック1300-1~1300-12）は、それらのトラック参照ボックス（1320-1~13

40

50

20 - 12) の値「h b a s」を有する。

【0191】

また、タイル化の通知は、トラックレベルで、サンプルレベルで、または、トラックおよびサンプルレベルであってもよい。

【0192】

デフォルトでは、タイルトラックは表示可能でないとみなされる点に留意する必要がある。しかし、MPEG-4 標準規格に準拠する高度なパーサーは、例えば、ストリーミングマニフェストファイルにおいて「t r e f」ボックスを見ることによって、表示可能なタイルトラックを検出して、それらを露出することができる、(タイルトラックが「h b a s」タイプの参照タイプを含む場合には、それは表示可能であるとみなすことができる)。これは、このタイルトラックが、ハンドラボックスで「タイル」値でマークされている場合であっても、標準ビデオトラックとみなすことができることを意味する。タイル化の通知がサンプルベースである場合には、タイル化情報がサンプルテーブルボックス(S a m p l e T a b l e B o x)として知られるボックスに配置されるので、タイルトラックまたは下位層トラックはそれらのハンドラボックスにおいて「v i d e」のタグを付けることができる。

【0193】

図14は、本発明の別の実施形態による、標準ビデオトラックとして再生可能な複合トラック、初期化データトラック、および独立タイルトラックを含む1組のトラックとしてHEVCビットストリームをカプセル化する例を示す。例示のために、カプセル化されたビットストリームは、図2に模式的に示したビデオシーケンスに対応する。

【0194】

図14に示すHEVCビットストリームのカプセル化は、主に、初期化データが専用の初期化データトラック1400(複合トラック1310ではない)に配置されるという点で、図13に示すものと異なる。

【0195】

図13を参照して説明したものと比較して、このような実施形態によって提供される利点の1つは、タイルトラックが独立に再生される場合に送信されるデータ量を対象とすることである。初期化データが専用のトラックで送信されるので、複合トラックを送信することは要求されない。

【0196】

HEVCファイルフォーマットの現行の仕様によれば、ファイルフォーマットのパラメータセット(P S)を送る2つの可能性が存在することを想起すべきであり、それはサンプルエン트리として知られるボックスにおいて、すなわちサンプルエントリボックスにおいて、およびデータサンプルにおいて、である。これらの2つの構成は、サンプルテーブルとして知られるボックスの「h v c 1」および「h e v 1」ボックスで、それぞれ通知される。サンプルのパラメータを記憶することはより複雑であるが、それはパラメータセットの更新の場合により多くのダイナミズムを可能にする。したがって、好ましい実施形態では、特にタイル化構成の変更のために、画像パラメータセット(P P S)の変更を扱うことを可能にするために、パラメータセットは、サンプルエントリボックスで、およびデータサンプルで(サンプルテーブルボックスのHEVCサンプルエントリパラメータの「h e v 1」値により)送られる。

【0197】

したがって、専用の初期化データトラック1400は、タイプが32、33、または34に等しい、NALユニットに類似の非VCL HEVC NALユニットだけをデータとして含み、これらのタイプは、ビデオパラメータセット、シーケンスパラメータセット、または画像パラメータセットにそれぞれ対応する。

【0198】

図14に示すように、タイルトラック1410-1~1410-12のメディアデータボックス「m d a t」の始めに位置するエクストラクタ1415-1~1415-12は

、専用の初期化データトラック 1400 のデータを示す。同様に、複合トラック 1405 の最初のエクストラクタ (1420) は、専用の初期化データトラック 1400 のデータを示す。したがって、初期化データトラック 1400 は、他のいかなるトラックも参照しないカプセル化された H E V C ビットストリームの唯一のトラックである。このように、初期化データトラック 1400 (i d = 2) に関する t r e f ボックスにおいて示される従属性がないので (t r e f ボックスには「 h b a s 」従属性がない)、初期化データトラック 1400 は独立に表示可能ではないとみなされる。

【 0 1 9 9 】

ビデオビットストリームにおいていくつかの初期化データが変更される場合には (すなわち画像パラメータセットが H E V C ビットストリームで生じる場合には)、符号 1425 で示すように、それらは変更が生じる時間的位置でサンプルデータに配置される。符号 1430 および符号 1435 - 1 ~ 1435 - 12 で示す、対応するエクストラクタは、複合トラック 1405 およびタイルトラック 1410 - 1 ~ 1410 - 12 の各々に、すなわちこれらの新しい P P S を参照する各タイルトラックにそれぞれ挿入される。

【 0 2 0 0 】

カプセル化された H E V C ビットストリームの各トラックにおいて、サンプル (および関係する N A L U) は、時間的順序で構成される。同様に、画像パラメータセットは、専用の初期化データトラック 1400 に時間的順序で構成される。「 t r u n 」ボックス (図 14 には示さず) は、正しい復号化時間を各サンプルに提供することを可能にする。

【 0 2 0 1 】

当然、局所のおよび特定の要件を満たすために、当業者は、多くの修正および変更を上記した解決策に適用することができるが、それらの全ては以下の特許請求の範囲によって規定される本発明の保護の範囲内に含まれる。

【 0 2 0 2 】

先行技術を鑑み、発明者は、タイル化時限メディアデータをサーバーにカプセル化するための、ならびに、複数のメディアセグメントファイルにカプセル化されたタイル化時限メディアデータから時限メディアデータ・ビットストリームを提供するための方法およびデバイスを提供する。

【 0 2 0 3 】

上述したように先行技術の欠点を改善することが、本発明の幅広い目的である。

【 0 2 0 4 】

本発明の第 1 の態様によれば、分割された時限メディアデータをサーバーにカプセル化する方法が提供され、分割された時限メディアデータは、時限サンプルを含み、各時限サンプルは、複数のサブサンプルを含み、本方法は、

時限サンプルのうちの 1 つの複数のサブサンプルの中から、少なくとも 1 つのサブサンプルを選択するステップと、

各選択されたサブサンプルについて、選択されたサブサンプルおよび他の時限サンプルの各々の 1 つの対応するサブサンプルを含む 1 つの分割トラックを生成するステップと、

生成された分割トラックのうちの少なくとも 1 つを識別する少なくとも 1 つのエクストラクタを含む少なくとも 1 つの参照トラックを生成するステップと、

生成されたトラックの各々を少なくとも 1 つのメディアセグメントファイルに独立にカプセル化するステップと、を含む。

【 0 2 0 5 】

したがって、本発明は、異なる部分、特に異なるタイルの組み合わせを可能にし、そして、どのような部分またはタイルの選択されたセットがクライアントデバイスによってダウンロードされても、パーシング時において有効なファイルフォーマットおよびエレメンタリストリームの作成を可能にする。

【 0 2 0 6 】

したがって、本発明の方法は、独立部分すなわちタイルの効率的なストリーミングに好適であり、有益なデータだけをクライアントデバイスに送信すればよく、2 つ以上の部分

10

20

30

40

50

すなわちタイルの選択（すなわち任意のROI）に好適であり、インデクセーションオーバーヘッド（バイト範囲要求と比較して）を低減し、MPEG標準規格に統合することができる。

【0207】

一実施形態では、分割された時限メディアデータはタイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも1つの分割トラックは、少なくとも1つのタイルトラックであり、少なくとも1つの参照トラックは、少なくとも1つの複合トラックである。少なくとも1つの参照トラックは、エクストラクタおよび選択的に初期化データを含んでもよい。

【0208】

一実施形態では、少なくとも1つのエクストラクタは、少なくとも1つの識別されたタイルトラックの少なくとも1つの空間サブサンプルをさらに識別する。

【0209】

一実施形態では、本方法は、生成されたタイルトラックおよび生成された少なくとも1つの複合トラックから、タイル化時限メディアデータの空間部分の記述を可能にするパラメータを含む初期化セグメントファイルを生成するステップをさらに含む。初期化セグメントファイルは、生成されたタイルトラックの参照、少なくとも1つの生成された複合トラックと関係する参照されたタイルトラックのリスト、トラックが空間サブサンプル情報を含むことを示す生成されたタイルトラックの各々と関係するタイルハンドラ、および/または、生成されたタイルトラックの各々と関係する一般プレゼンテーション情報を含んでもよい。

【0210】

一実施形態では、各選択された空間サブサンプルについて1つのタイルトラックを生成するステップにおいて、少なくとも2つのタイルトラックが生成され、少なくとも1つの複合トラックは、同じ時限サンプル内に含まれる2つの空間サブサンプルのいずれかを復号化するために用いられるデータの少なくとも1つの項目を含む。

【0211】

一実施形態では、タイル化時限メディアデータは、時限サンプルの各々が参照層および少なくとも1つの拡張層を含むことにより、スケーラブルなタイル化時限メディアデータであり、参照層は少なくとも1つの参照空間サブサンプルを含み、少なくとも1つの拡張層は複数の拡張空間サブサンプルを含み、各選択された空間サブサンプルについて1つのタイルトラックを生成するステップにおいて生成されるタイルトラックは、拡張空間サブサンプルを含む拡張タイルトラックであり、本方法は、参照空間サブサンプルを含む少なくとも1つの参照トラックを生成するステップをさらに含む。

【0212】

一実施形態では、参照層は、複数の参照空間サブサンプルを含み、複数の参照トラックが生成され、複数の参照トラックの各参照トラックは、参照タイルトラックを形成する。

【0213】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、対応するトラックの定義を含むメタデータを記憶するためのデータ構造を含む。

【0214】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルのうちの少なくとも1つは、対応するトラックのサブサンプルの少なくとも1つのグループの少なくとも1つの定義を含むメタデータを記憶するためのデータ構造を含む。

【0215】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルのうちの少なくとも1つは、対応するトラックのサブサンプルの種々のグループの種々の定義を含むメタデータを記憶するためのデータ構造を含む。

10

20

30

40

50

【0216】

一実施形態では、メタデータを記憶するデータ構造は、少なくとも1つの定義、および少なくとも1つの定義が対応するトラックの全てのサブサンプルに当てはまることを示すバージョンタイプを含む。

【0217】

一実施形態では、対応するトラックのサブサンプルの少なくとも1つのグループの少なくとも1つの定義を含むメタデータを記憶するデータ構造の少なくとも1つのパラメータは、サブサンプルの復号化従属性に関して、グループの各サブサンプルが別のグループのサブサンプルを用いずに復号化され得ることを示す。

【0218】

一実施形態では、グループのサブサンプルは、サブサンプルが属するグルーピングタイプに従って識別される。

【0219】

一実施形態では、グループのサブサンプルは、サブサンプルが属するグルーピングタイプに従って識別され、グルーピングタイプは、グループの各サブサンプルについての定義を提供する。

【0220】

一実施形態では、初期化セグメントファイルは、対応する少なくとも1つのタイルトラックをカプセル化することにより生じるメディアセグメントファイルに記憶された任意の特定の定義に関係しない、少なくとも1つのタイルトラックのサブサンプルを処理するために、デフォルトで用いられる1つの定義を含む。

【0221】

一実施形態では、本方法は、デフォルトで用いられる初期化セグメントファイルに記憶された定義を変更することを可能にするために、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルの少なくとも1つに初期化データを追加するステップをさらに含む。

【0222】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、空間サブサンプルを記憶するためのデータ構造を含む。

【0223】

生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、エクストラクタを記憶するためのデータ構造を含む。

【0224】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルの少なくとも1つは、メディアセグメントファイルの少なくとも1つが標準メディアセグメントファイルとして独立に処理され得るように、参照トラックに記憶された初期化データを含む。

【0225】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルの少なくとも1つは、参照トラックに記憶された初期化データを識別する少なくとも1つのエクストラクタを記憶するためのデータ構造を含む。

【0226】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルの少なくとも1つは、メディアセグメントファイルの少なくとも1つが標準メディアセグメントファイルとして独立に処理され得るように、複合トラックの初期化データを識別する少なくとも1つのエクストラクタを記憶するためのデータ構造を含む。

【0227】

一実施形態では、本方法は、メディアセグメントファイルの少なくとも1つが標準メディアセグメントファイルとして独立に処理され得ることを示すフラグをセットするステッ

10

20

30

40

50

プをさらに含む。

【0228】

一実施形態では、本方法は、初期化データを含む少なくとも1つの初期化データトラックを生成するステップをさらに含み、少なくとも1つの初期化データトラックは、少なくとも1つのメディアセグメントファイルに独立にカプセル化される。

【0229】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルの少なくとも1つは、メディアセグメントファイルの少なくとも1つが標準メディアセグメントファイルとして独立に処理され得るように、初期化データトラックに記憶された初期化データを識別する少なくとも1つのエクストラクタを記憶するためのデータ構造を含む。

10

【0230】

一実施形態では、本方法は、メディアセグメントファイルの少なくとも1つが標準メディアセグメントファイルとして独立に処理され得ることを示すフラグをセットするステップをさらに含む

一実施形態では、サーバーは、ハイパーテキスト・トランスファー・プロトコル（HTTP）と互換性を有する。

【0231】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、国際標準化機構によって定義されたベース・メディア・ファイル・フォーマットおよびHTTPフォーマット上のダイナミック・アダプティブ・ストリーミングと互換性を有する。

20

【0232】

一実施形態では、分割された時限メディアデータをサーバーにカプセル化する方法であって、分割された時限メディアデータは、時限サンプルを含み、各時限サンプルは、複数の空間サブサンプルを含み、本方法は、

時限サンプルのうちの1つの複数のサブサンプルの中から、少なくとも1つのサブサンプルを選択するステップと、

各選択されたサブサンプルについて、選択されたサブサンプルおよび他の時限サンプルの各々の1つの対応するサブサンプルを含む1つの分割トラックを生成するステップと、

30

第1の分割トラックに含まれる第1の空間サブサンプルおよび第2の分割トラックに含まれる第2の空間サブサンプルを関連づける情報を含む、少なくとも1つの参照トラックを生成するステップと、

各生成されたトラックを少なくとも1つのメディアセグメントファイルに独立にカプセル化するステップと、を含む。

【0233】

このような実施形態は、本発明の第1の態様を参照することにより、上で開示した上述の実施形態の1つまたはいくつかと組み合わせることができる。

【0234】

本発明の第2態様は、複数のメディアセグメントファイルにカプセル化された分割された時限メディアデータから時限メディアデータ・ビットストリームを提供する方法を提供し、クライアントデバイスにおいて、分割された時限メディアデータは時限サンプルを含み、各時限サンプルは複数のサブサンプルを含み、複数のメディアセグメントファイルは、少なくとも1つの参照トラックおよび少なくとも1つの分割トラックを含み、少なくとも1つの分割トラックの各々は、時限サンプルのうちの1つの複数のサブサンプルの中から選択された1つのサブサンプルを含み、かつ、他の時限サンプルの各々の1つの対応する空間サブサンプルを含み、少なくとも1つの参照トラックは、少なくとも1つの分割トラックを識別し、かつ、識別された分割トラックの少なくとも1つのサブサンプルを識別する、少なくとも1つのエクストラクタを含み、本方法は、

40

サブサンプルを表す情報の項目を選択するステップと、

50

情報の選択された項目に対応する選択されたサブサンプルを含む少なくとも1つの分割トラックを要求するステップと、

複数のメディアセグメントファイルを受け取るステップと、

時限メディアデータ・ビットストリームを生成するために、少なくとも1つの参照トラックのエクストラクタを、対応する識別されたサブサンプルと置き換えるステップと、を含む。

【0235】

したがって、本発明は、異なる部分、特に異なるタイルの組み合わせを可能にし、そして、どのような部分またはタイルの選択されたセットがクライアントデバイスによってダウンロードされても、パーシング時において有効なファイルフォーマットおよびエレメンタリストリームの作成を可能にする。

10

【0236】

したがって、本発明の方法は、独立部分すなわちタイルの効率的なストリーミングに好適であり、有益なデータだけをクライアントデバイスに送信すればよく、2つ以上の部分すなわちタイルの選択（すなわち任意のROI）に好適であり、インデクセーションオーバーヘッド（バイト範囲要求と比較して）を低減し、MPEG標準規格に統合することができる。

【0237】

一実施形態では、分割された時限メディアデータはタイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも1つの分割トラックは、少なくとも1つのタイルトラックであり、少なくとも1つの参照トラックは、少なくとも1つの複

20

合トラックである。

【0238】

一実施形態では、分割された時限メディアデータはタイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも1つの分割トラックは、少なくとも1つのタイルトラックであり、少なくとも1つの参照トラックは、エクストラクタおよび選択的に初期化データを含む少なくとも1つの複合トラックである。

【0239】

一実施形態では、本方法は、

少なくとも1つの複合トラックのエクストラクタからトラック参照を取得するステップと、

30

取得されたトラック参照に対応するトラックが受け取られたか否かをチェックするステップと、

取得されたトラック参照に対応するトラックが受け取られなかった場合には、時限メディアデータ・ビットストリームを生成するために対応するエクストラクタを除去するステップと、をさらに含む。

【0240】

一実施形態では、本方法は、取得されたトラック参照に対応するトラックがタイルタイプのトラックであることを、検証するステップをさらに含む。

【0241】

40

一実施形態では、本方法は、

少なくとも1つの複合トラックのエクストラクタからトラック参照を取得するステップと、

取得されたトラック参照に対応するトラックが受け取られたか否かをチェックするステップと、

取得されたトラック参照に対応するトラックが受け取られなかった場合には、時限メディアデータ・ビットストリームを生成するために、対応するエクストラクタをパディングに置き換えるステップと、をさらに含む。

【0242】

一実施形態では、本方法は、取得されたトラック参照に対応するトラックがタイルタイ

50

プのトラックであることを、検証するステップをさらに含む。

【 0 2 4 3 】

一実施形態では、本方法は、生成されたタイルトラックおよび生成された少なくとも 1 つの複合トラックから、タイル化時限メディアデータの空間部分の形成を可能にするパラメータを含む初期化セグメントファイルを受け取るステップをさらに含む。

【 0 2 4 4 】

一実施形態では、少なくとも 2 つのタイルトラックを表す少なくとも 2 つのメディアセグメントファイルが受け取られ、少なくとも 1 つの複合トラックは、空間サブサンプルの同じセットに属する 2 つの空間サブサンプルのいずれかを復号化するために用いられるデータの少なくとも 1 つの項目を含む。

10

【 0 2 4 5 】

一実施形態では、タイル化時限メディアデータは、時限サンプルの各々が参照層および少なくとも 1 つの拡張層を含むことにより、スケーラブルなタイル化時限メディアデータであり、参照層は少なくとも 1 つの参照空間サブサンプルを含み、少なくとも 1 つの拡張層は複数の拡張空間サブサンプルを含み、少なくとも 2 つのタイルトラックを表す少なくとも 2 つのメディアセグメントファイルが受け取られ、少なくとも 2 つの受け取ったトラックの一方は、参照空間サブサンプルを含む参照トラックであり、少なくとも 2 つの受け取ったトラックの他方は、拡張空間サブサンプルを含む拡張タイルトラックである。

【 0 2 4 6 】

一実施形態では、参照層は、複数の参照空間サブサンプルを含み、各々が参照タイルトラックを形成する複数の参照トラックが受け取られる。

20

【 0 2 4 7 】

一実施形態では、メディアセグメントファイルの少なくとも 1 つは、対応するトラックのサブサンプルの少なくとも 1 つのグループの少なくとも 1 つの定義を含むメタデータを記憶するためのデータ構造を含み、本方法は、対応するトラックのサブサンプルの少なくとも 1 つのグループを処理するための少なくとも 1 つの定義を取得するステップを含む。

【 0 2 4 8 】

一実施形態では、メディアセグメントファイルの少なくとも 1 つは、参照トラックに記憶された初期化データを含み、本方法は、メディアセグメントファイルの少なくとも 1 つを標準メディアセグメントファイルとして独立に処理するステップをさらに含む。

30

【 0 2 4 9 】

一実施形態では、メディアセグメントファイルの少なくとも 1 つは、参照トラックに記憶された初期化データを識別する少なくとも 1 つのエクストラクタを記憶するためのデータ構造を含み、本方法は、メディアセグメントファイルの少なくとも 1 つを処理する場合に、参照トラックに記憶されたデータにアクセスするステップをさらに含む。

【 0 2 5 0 】

一実施形態では、メディアセグメントファイルの少なくとも 1 つは、複合トラックの初期化データを識別する少なくとも 1 つのエクストラクタを記憶するためのデータ構造を含み、本方法は、メディアセグメントファイルの少なくとも 1 つを標準メディアセグメントファイルとして独立に処理するために、参照トラックに記憶されたデータにアクセスするステップをさらに含む。

40

【 0 2 5 1 】

一実施形態では、本方法は、初期化データを含む少なくとも 1 つの初期化データトラックを取得するステップをさらに含む、少なくとも 1 つの初期化データトラックは、少なくとも 1 つのメディアセグメントファイルに独立にカプセル化される。

【 0 2 5 2 】

一実施形態では、メディアセグメントファイルの少なくとも 1 つは、初期化データトラックに記憶された初期化データを識別する少なくとも 1 つのエクストラクタを記憶するためのデータ構造を含み、本方法は、メディアセグメントファイルの少なくとも 1 つを独立に処理するために、参照トラックに記憶されたデータにアクセスするステップをさらに含

50

む。

【 0 2 5 3 】

一実施形態では、クライアントデバイスは、ハイパーテキスト・トランスファー・プロトコル (H T T P) と互換性を有する。

【 0 2 5 4 】

一実施形態では、受け取ったメディアセグメントファイルは、国際標準化機構によって定義されたベース・メディア・ファイル・フォーマットおよび H T T P フォーマット上のダイナミック・アダプティブ・ストリーミングと互換性を有する。

【 0 2 5 5 】

本発明の第 3 の態様によれば、上記の方法の各ステップを実行するように適応される手段を含むデバイスが提供される。

10

【 0 2 5 6 】

本発明の第 4 の態様によれば、分割された時限メディアデータをサーバーにカプセル化するためのデバイスが提供され、分割された時限メディアデータは時限サンプルを含み、各時限サンプルは複数のサブサンプルを含み、デバイスは、

時限サンプルのうちの 1 つの複数のサブサンプルの中から、少なくとも 1 つのサブサンプルを選択するステップと、

各選択されたサブサンプルについて、選択されたサブサンプルおよび他の時限サンプルの各々の 1 つの対応するサブサンプルを含む 1 つの分割トラックを生成するステップと、

生成された分割トラックのうちの少なくとも 1 つを識別する少なくとも 1 つのエクストラクタを含む少なくとも 1 つの参照トラックを生成するステップと、

生成されたトラックの各々を少なくとも 1 つのメディアセグメントファイルに独立にカプセル化するステップと、を実行するように構成される少なくとも 1 つのマイクロプロセッサを含む。

20

【 0 2 5 7 】

したがって、本発明は、異なる部分、特に異なるタイルの組み合わせを可能にし、そして、どのような部分またはタイルの選択されたセットがクライアントデバイスによってダウンロードされても、パーシング時において有効なファイルフォーマットおよびエレメンタリストリームの作成を可能にする。

【 0 2 5 8 】

したがって、本発明の方法は、独立部分すなわちタイルの効率的なストリーミングに好適であり、有益なデータだけをクライアントデバイスに送信すればよく、2 つ以上の部分すなわちタイルの選択 (すなわち任意の R O I) に好適であり、インデクセーションオーバーヘッド (バイト範囲要求と比較して) を低減し、M P E G 標準規格に統合することができる。

30

【 0 2 5 9 】

一実施形態では、分割された時限メディアデータはタイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも 1 つの分割トラックは、少なくとも 1 つのタイルトラックであり、少なくとも 1 つの参照トラックは、少なくとも 1 つの複合トラックである。

40

【 0 2 6 0 】

一実施形態では、分割された時限メディアデータは、タイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも 1 つの分割トラックは、少なくとも 1 つのタイルトラックであり、少なくとも 1 つの参照トラックは、エクストラクタおよび選択的に初期化データを含む少なくとも 1 つの複合トラックである。

【 0 2 6 1 】

一実施形態では、マイクロプロセッサは、生成されたタイルトラックおよび生成された少なくとも 1 つの複合トラックから、タイル化時限メディアデータの空間部分の形成を可能にするパラメータを含む初期化セグメントファイルを生成するステップを実行するようにさらに構成され、初期化セグメントファイルは、生成されたタイルトラックの参照を含

50

む。

【0262】

一実施形態では、初期化セグメントファイルは、少なくとも1つの生成された複合トラックと関係する参照されたタイルトラックのリストをさらに含む。

【0263】

一実施形態では、初期化セグメントファイルは、トラックが空間サブサンプル情報を含むことを示す生成されたタイルトラックの各々と関係するタイルハンドラをさらに含む。

【0264】

一実施形態では、初期化セグメントファイルは、生成されたタイルトラックの各々と関係する一般プレゼンテーション情報をさらに含む。

【0265】

一実施形態では、マイクロプロセッサは、各選択された空間サブサンプルについて1つのタイルトラックを生成するステップにおいて、少なくとも2つのタイルトラックが生成されるようにさらに構成され、少なくとも1つの複合トラックは、同じ時限サンプル内に含まれる2つの空間サブサンプルのいずれかを復号化するために用いられるデータの少なくとも1つの項目を含む。

【0266】

一実施形態では、タイル化時限メディアデータは、時限サンプルの各々が参照層および少なくとも1つの拡張層を含むことにより、スケーラブルなタイル化時限メディアデータであり、参照層は少なくとも1つの参照空間サブサンプルを含み、少なくとも1つの拡張層は複数の拡張空間サブサンプルを含み、マイクロプロセッサは、各選択された空間サブサンプルについて1つのタイルトラックを生成するステップにおいて生成されるタイルトラックが拡張空間サブサンプルを含む拡張タイルトラックであるように、さらに構成され、マイクロプロセッサは、参照空間サブサンプルを含む少なくとも1つの参照トラックを生成するステップを実行するようにさらに構成される。

【0267】

一実施形態では、参照層は、複数の参照空間サブサンプルを含み、マイクロプロセッサは、複数の参照トラックが生成されるようにさらに構成され、複数の参照トラックの各参照トラックは、参照タイルトラックを形成する。

【0268】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、対応するトラックの定義を含むメタデータを記憶するためのデータ構造を含む。

【0269】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、空間サブサンプルを記憶するためのデータ構造を含む。

【0270】

一実施形態では、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、エクストラクタを記憶するためのデータ構造を含む。

【0271】

一実施形態では、サーバーは、ハイパーテキスト・トランスファー・プロトコル（HTTP）と互換性を有し、生成されたトラックの各々をカプセル化することにより生じるメディアセグメントファイルは、国際標準化機構によって定義されたベース・メディア・ファイル・フォーマットおよびHTTPフォーマット上のダイナミック・アダプティブ・ストリーミングと互換性を有する。

【0272】

一実施形態では、分割された時限メディアデータをサーバーにカプセル化するためのデバイスであって、分割された時限メディアデータは時限サンプルを含み、各時限サンプルは複数の空間サブサンプルを含み、デバイスは、

時限サンプルのうちの1つの複数のサブサンプルの中から、少なくとも1つのサブサン

10

20

30

40

50

ブルを選択するように構成される選択部と、

各選択された空間サブサンプルについて、選択されたサブサンプルおよび他の時限サンプルの各々の1つの対応するサブサンプルを含む1つの分割トラックを生成するように構成される第1の生成部と、

第1の分割トラックに含まれる第1の空間サブサンプルおよび第2の分割トラックに含まれる第2の空間サブサンプルを関連づける情報を含む、少なくとも1つの参照トラックを生成するように構成される第2の生成部と、

各生成されたトラックを少なくとも1つのメディアセグメントファイルに独立にカプセル化するように構成されるカプセル化部と、を含む。

【0273】

このような実施形態は、本発明の第4の態様を参照することにより、上で開示した上述の実施形態の1つまたはいくつかと組み合わせることができる。

【0274】

本発明の第5の態様によれば、上記のデバイスを含むビデオエンコーダが提供される。

【0275】

本発明の第6の態様によれば、複数のメディアセグメントファイルにカプセル化された分割された時限メディアデータから時限メディアデータ・ビットストリームを提供するためのデバイスが提供され、クライアントデバイスにおいて、分割された時限メディアデータは時限サンプルを含み、各時限サンプルは複数のサブサンプルを含み、複数のメディアセグメントファイルは、少なくとも1つの参照トラックおよび少なくとも1つの分割トラックを含み、少なくとも1つの分割トラックの各々は、時限サンプルのうちの1つの複数のサブサンプルの中から選択された1つのサブサンプルを含み、かつ、他の時限サンプルの各々の1つの対応するサブサンプルを含み、少なくとも1つの参照トラックは、少なくとも1つの分割トラックを識別し、かつ、識別された分割トラックの少なくとも1つのサブサンプルを識別する、少なくとも1つのエクストラクタを含み、デバイスは、

サブサンプルを表す情報の項目を選択するステップと、

情報の選択された項目に対応する選択されたサブサンプルを含む少なくとも1つの分割トラックを要求するステップと、

複数のメディアセグメントファイルを受け取るステップと、

時限メディアデータ・ビットストリームを生成するために、少なくとも1つの参照トラックのエクストラクタを、対応する識別されたサブサンプルと置き換えるステップと、を実行するように構成される少なくとも1つのマイクロプロセッサを含む。

【0276】

したがって、本発明は、異なる部分、特に異なるタイルの組み合わせを可能にし、そして、どのような部分またはタイルの選択されたセットがクライアントデバイスによってダウンロードされても、パーシング時において有効なファイルフォーマットおよびエレメンタリストリームの作成を可能にする。

【0277】

したがって、本発明の方法は、独立部分すなわちタイルの効率的なストリーミングに好適であり、有益なデータだけをクライアントデバイスに送信すればよく、2つ以上の部分すなわちタイルの選択（すなわち任意のROI）に好適であり、インデクセーションオーバーヘッド（バイト範囲要求と比較して）を低減し、MPEG標準規格に統合することができる。

【0278】

一実施形態では、分割された時限メディアデータはタイル化時限メディアデータであり、サブサンプルは空間サブサンプルであり、少なくとも1つの分割トラックは、少なくとも1つのタイルトラックであり、少なくとも1つの参照トラックは、少なくとも1つの複合トラックである。

【0279】

一実施形態では、分割された時限メディアデータは、タイル化時限メディアデータであ

10

20

30

40

50

り、サブサンプルは空間サブサンプルであり、少なくとも1つの分割トラックは、少なくとも1つのタイルトラックであり、少なくとも1つの参照トラックは、エクストラクタおよび選択的に初期化データを含む少なくとも1つの複合トラックである。

【0280】

一実施形態では、マイクロプロセッサは、
少なくとも1つの複合トラックのエクストラクタからトラック参照を取得するステップと、

取得されたトラック参照に対応するトラックが受け取られたか否かをチェックするステップと、

取得されたトラック参照に対応するトラックが受け取られなかった場合には、時限メディアデータ・ビットストリームを生成するに対応するエクストラクタを除去するステップと、を実行するようにさらに構成される。

10

【0281】

一実施形態では、マイクロプロセッサは、取得されたトラック参照に対応するトラックがタイルタイプのトラックであることを、検証するステップを実行するようにさらに構成される。

【0282】

一実施形態では、マイクロプロセッサは、
少なくとも1つの複合トラックのエクストラクタからトラック参照を取得するステップと、

20

取得されたトラック参照に対応するトラックが受け取られたか否かをチェックするステップと、

取得されたトラック参照に対応するトラックが受け取られなかった場合には、時限メディアデータ・ビットストリームを生成するために、対応するエクストラクタをパディングに置き換えるステップと、を実行するようにさらに構成される。

【0283】

一実施形態では、マイクロプロセッサは、取得されたトラック参照に対応するトラックがタイルタイプのトラックであることを、検証するステップを実行するようにさらに構成される。

【0284】

30

一実施形態では、マイクロプロセッサは、生成されたタイルトラックおよび生成された少なくとも1つの複合トラックから、タイル化時限メディアデータの空間部分の形成を可能にするパラメータを含む初期化セグメントファイルを受け取るステップを実行するようにさらに構成される。

【0285】

一実施形態では、マイクロプロセッサは、少なくとも2つのタイルトラックを表す少なくとも2つのメディアセグメントファイルが受け取られるようにさらに構成され、少なくとも1つの複合トラックは、空間サブサンプルの同じセットに属する2つの空間サブサンプルのいずれかを復号化するために用いられるデータの少なくとも1つの項目を含む。

【0286】

40

一実施形態では、タイル化時限メディアデータは、時限サンプルの各々が参照層および少なくとも1つの拡張層を含むことにより、スケーラブルなタイル化時限メディアデータであり、参照層は少なくとも1つの参照空間サブサンプルを含み、少なくとも1つの拡張層は複数の拡張空間サブサンプルを含み、マイクロプロセッサは、少なくとも2つのタイルトラックを表す少なくとも2つのメディアセグメントファイルが受け取られるようにさらに構成され、少なくとも2つの受け取ったトラックの一方は、参照空間サブサンプルを含む参照トラックであり、少なくとも2つの受け取ったトラックの他方は、拡張空間サブサンプルを含む拡張タイルトラックである。

【0287】

一実施形態では、クライアントデバイスは、ハイパーテキスト・トランスファー・プロ

50

トコル（ＨＴＴＰ）と互換性を有し、受け取ったメディアセグメントファイルは、国際標準化機構によって定義されたベース・メディア・ファイル・フォーマットおよびＨＴＴＰフォーマット上のダイナミック・アダプティブ・ストリーミングと互換性を有する。

【０２８８】

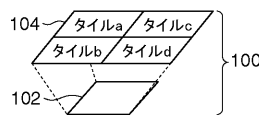
本発明の第７の態様によれば、上記のデバイスを含むビデオデコーダが提供される。

【０２８９】

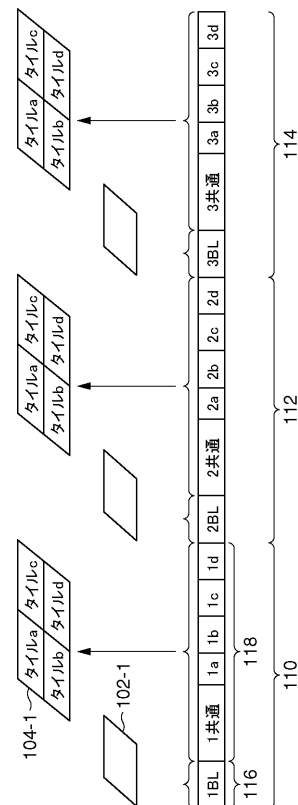
本発明はソフトウェアで実現することができるので、本発明は、任意の好適なキャリア媒体上の、プログラム可能な装置に供給するためのコンピュータ可読コードとして具現化することができる。有形のキャリア媒体は、フロッピーディスク、ＣＤ－ＲＯＭ、ハードディスクドライブ、磁気テープデバイス、または固体状態メモリデバイスなどの記憶媒体を含むことができる。一時的なキャリア媒体は、電気信号、電子信号、光信号、音響信号、磁気信号、または電磁信号、例えばマイクロ波もしくはＲＦ信号などの信号を含むことができる。

10

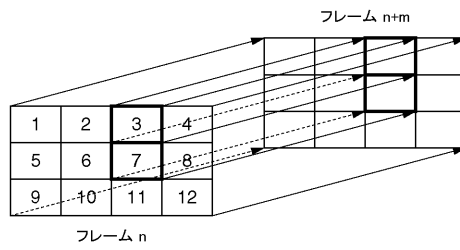
【図１ａ】



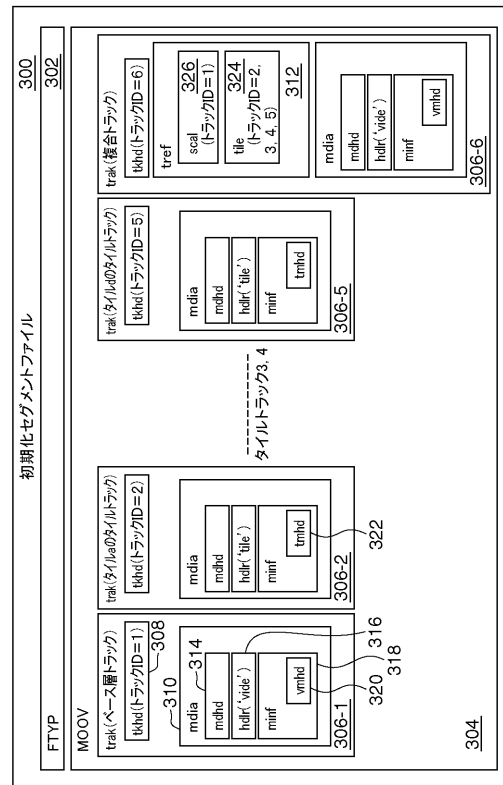
【図１ｂ】



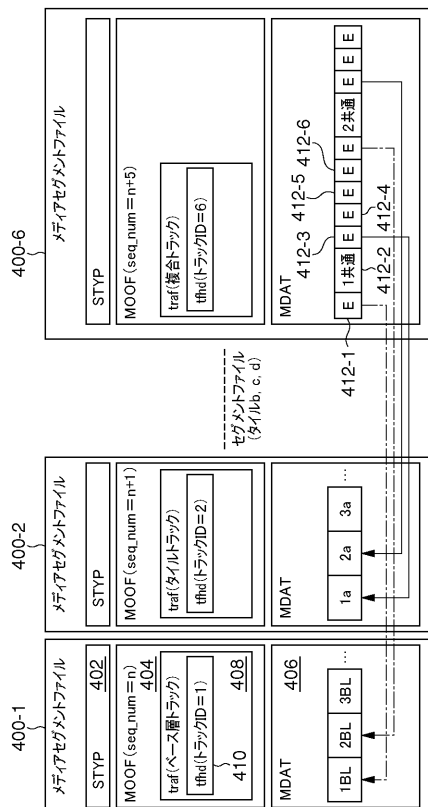
【 図 2 】



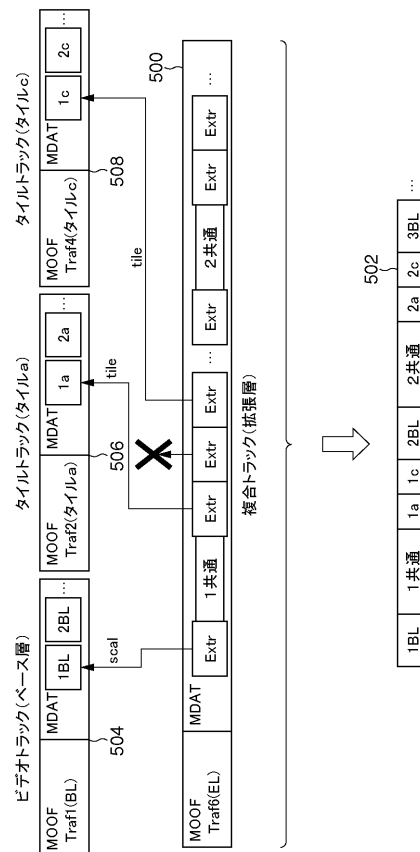
【 図 3 】



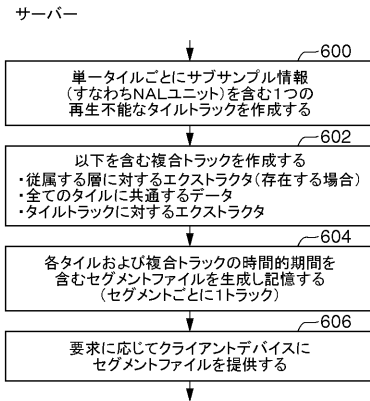
【圖 4】



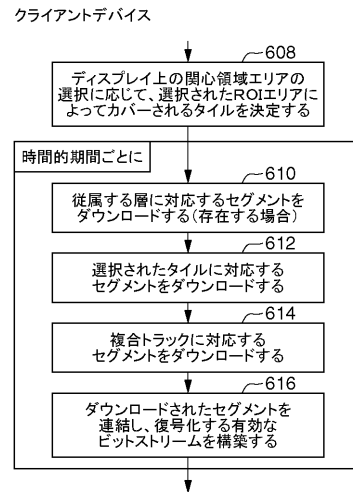
【 図 5 】



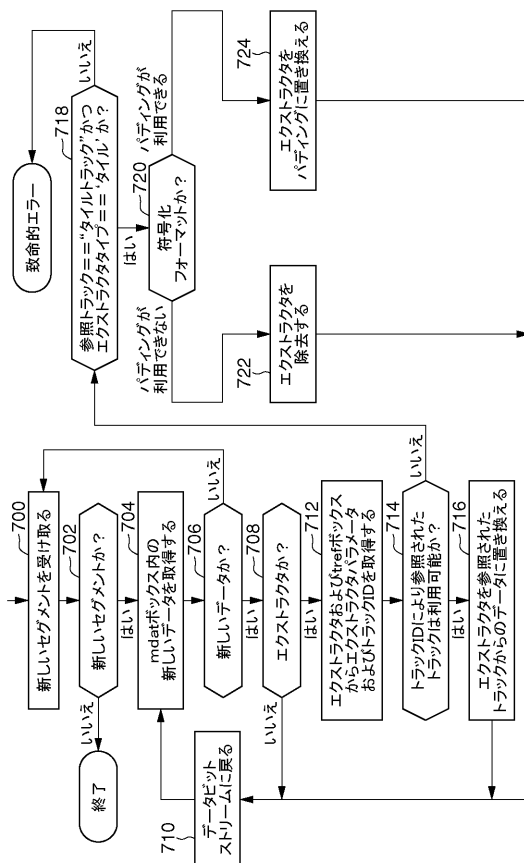
【図 6 a】



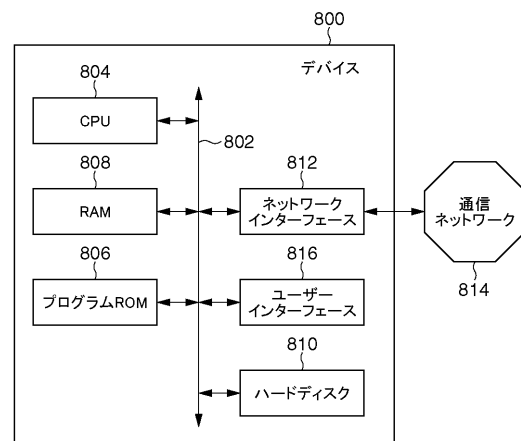
【図 6 b】



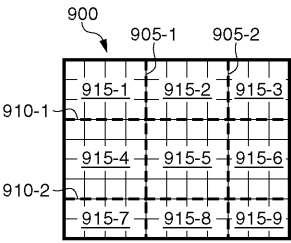
【図 7】



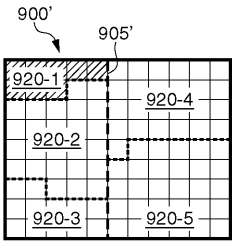
【図 8】



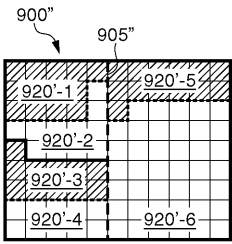
【図 9 a】



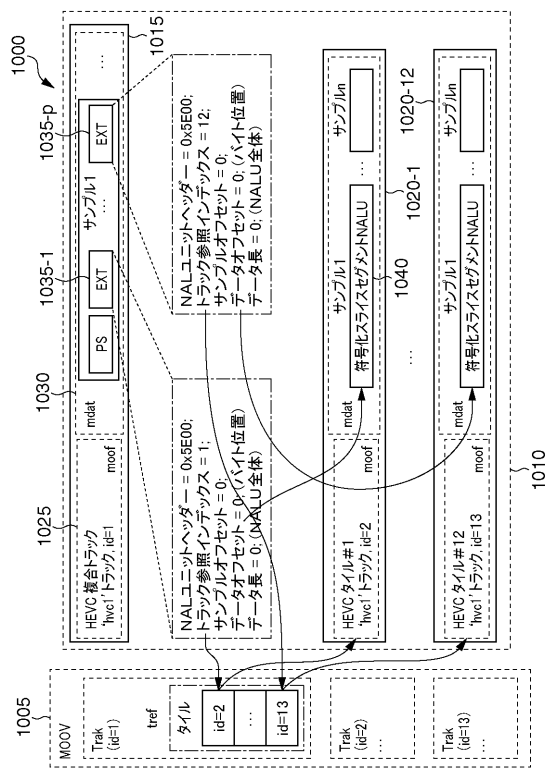
【図 9 b】



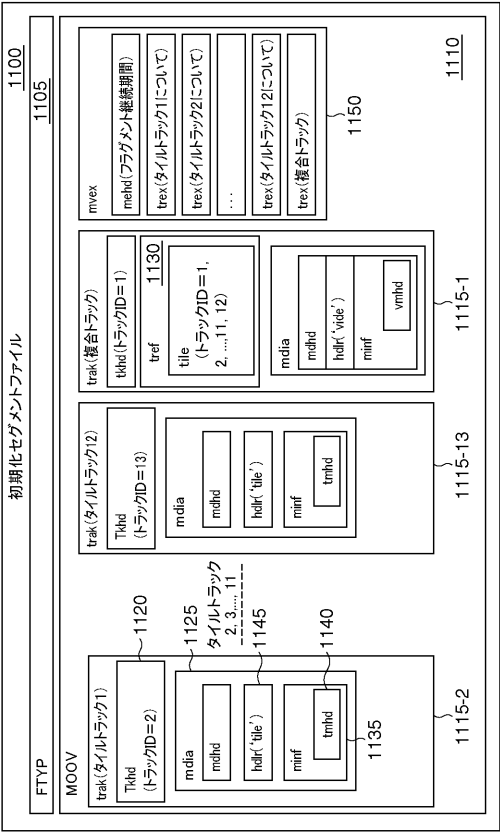
【図 9 c】



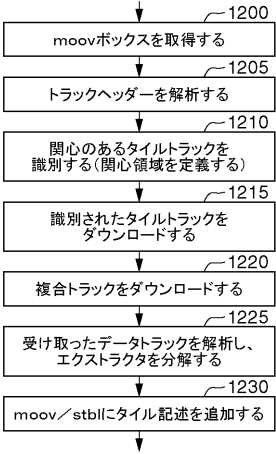
【図 10】



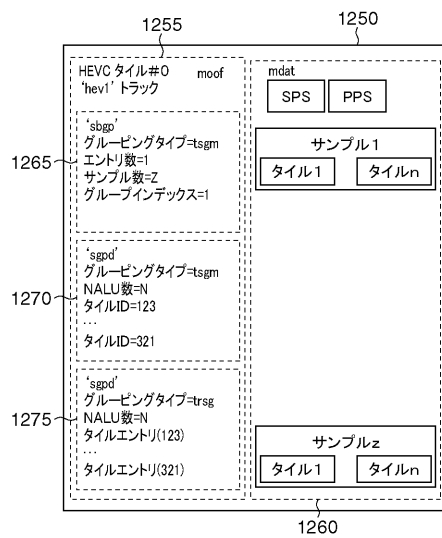
【図 11】



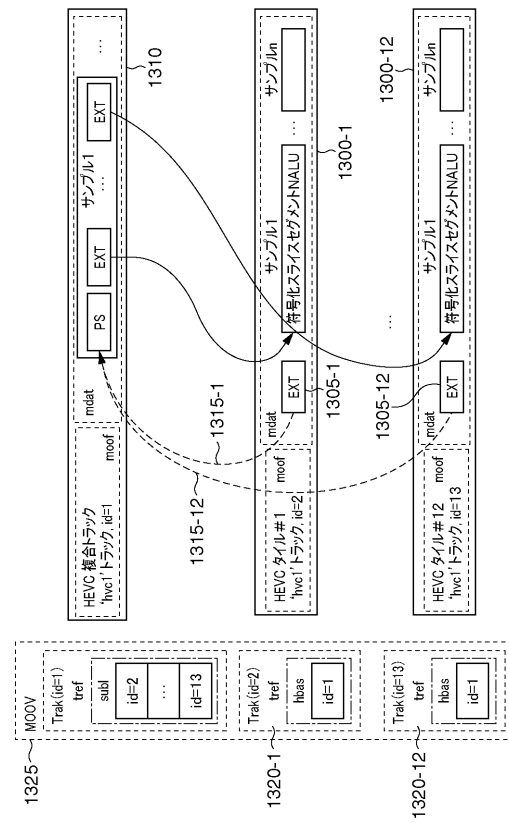
【図 12 a】



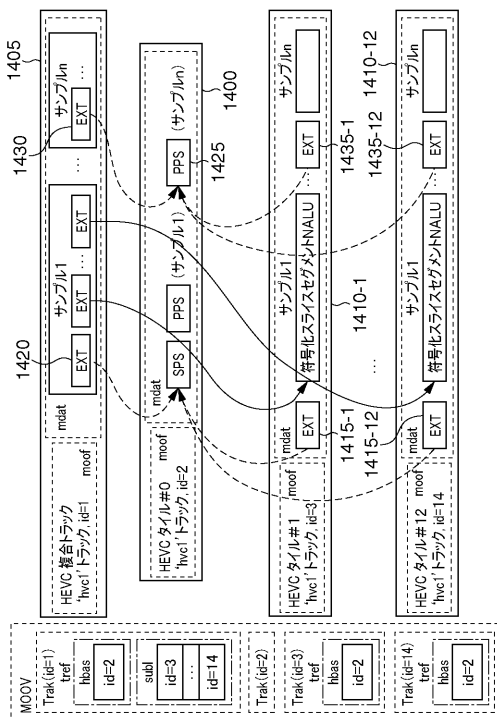
【図 12b】



【図 13】



【図 14】



フロントページの続き

(31)優先権主張番号 1306451.4

(32)優先日 平成25年4月9日(2013.4.9)

(33)優先権主張国・地域又は機関
英国(GB)

(72)発明者 エルヴェ ル フロッシュ

フランス国 レヌ - アタラント, セデックス セッソン - セヴィニエ 35517, リュ ド
ウ ラ トゥッシュ - ランベール キヤノン リサーチセンター フランス エス. エー. エ
ス. 内

(72)発明者 フランク ドゥヌアル

フランス国 レヌ - アタラント, セデックス セッソン - セヴィニエ 35517, リュ ド
ウ ラ トゥッシュ - ランベール キヤノン リサーチセンター フランス エス. エー. エ
ス. 内

(72)発明者 シリル コンコラト

フランス国 46 リュ パロー 75013 パリ テレコムパリステック内

(72)発明者 ジャン ルフェーブル

フランス国 46 リュ パロー 75013 パリ テレコムパリステック内

合議体

審判長 清水 正一

審判官 渡辺 努

審判官 檜本 剛

(56)参考文献 特開2003-224846(JP, A)

特開2009-164936(JP, A)

特表2010-531590(JP, A)

特表2009-503950(JP, A)

特表2005-524128(JP, A)

特表2011-528868(JP, A)

特開2011-87103(JP, A)

国際公開第2012/90501(WO, A1)

国際公開第2012/168365(WO, A1)

Ingo Kofler et al: "Implications of the ISO
base media file format on adaptive HTTP str
eaming of H.264/SVC", 2012 IEEE Consumer Com
munications and Networking Conference(CNC)
, IEEE, 2012年1月14日, p. 549 - 553

(58)調査した分野(Int.Cl., DB名)

H04N21/00-21/858

H04N19/00-19/98